agriculture

Special Issue Reprint

Intelligent Agricultural Machinery and Robots

Embracing Technological Advancements for a Sustainable and Highly Efficient Agricultural Future

Edited by
Jin Yuan and Zichen Huang

mdpi.com/journal/agriculture

MDPI

Intelligent Agricultural Machinery and Robots: Embracing Technological Advancements for a Sustainable and Highly Efficient Agricultural Future

Intelligent Agricultural Machinery and Robots: Embracing Technological Advancements for a Sustainable and Highly Efficient Agricultural Future

Guest Editors

Jin Yuan
Zichen Huang

MDPI

Basel • Beijing • Wuhan • Barcelona • Belgrade • Novi Sad • Cluj • Manchester

Guest Editors
Jin Yuan
School of Mechanical and Electronic Engineering
Shandong Agriculture University
Taian
China

Zichen Huang
College of Biosystems Engineering and Food Science
Zhejiang University
Hangzhou
China

Editorial Office
MDPI AG
Grosspeteranlage 5
4052 Basel, Switzerland

This is a reprint of the Special Issue, published open access by the journal *Agriculture* (ISSN 2077-0472), freely accessible at: https://www.mdpi.com/si/agriculture/XHR69P755G.

For citation purposes, cite each article independently as indicated on the article page online and as indicated below:

Lastname, A.A.; Lastname, B.B. Article Title. *Journal Name* **Year**, *Volume Number*, Page Range.

ISBN 978-3-7258-3211-8 (Hbk)
ISBN 978-3-7258-3212-5 (PDF)
https://doi.org/10.3390/books978-3-7258-3212-5

© 2025 by the authors. Articles in this book are Open Access and distributed under the Creative Commons Attribution (CC BY) license. The book as a whole is distributed by MDPI under the terms and conditions of the Creative Commons Attribution-NonCommercial-NoDerivs (CC BY-NC-ND) license (https://creativecommons.org/licenses/by-nc-nd/4.0/).

Contents

About the Editors . . . . . **vii**

Preface . . . . . **ix**

Jin Yuan and Zichen Huang
Intelligent Agricultural Machinery and Robots: Embracing Technological Advancements for a Sustainable and Highly Efficient Agricultural Future
Reprinted from: *Agriculture* **2024**, *14*, 2150, https://doi.org/10.3390/agriculture14122150 . . . . **1**

Chenbo Shi, Yuejia Li, Xin Jiang, Wenxin Sun, Changsheng Zhu, Yuanzheng Mo, et al.
Real-Time ConvNext-Based U-Net with Feature Infusion for Egg Microcrack Detection
Reprinted from: *Agriculture* **2024**, *14*, 1655, https://doi.org/10.3390/agriculture14091655 . . . . **4**

Shuzhen Yang, Jingmin Zhang and Jin Yuan
A High-Accuracy Contour Segmentation and Reconstruction of a Dense Cluster of Mushrooms Based on Improved SOLOv2
Reprinted from: *Agriculture* **2024**, *14*, 1646, https://doi.org/10.3390/agriculture14091646 . . . . **23**

Meng Wang, Changhe Niu, Zifan Wang, Yongxin Jiang, Jianming Jian and Xiuying Tang
Model and Parameter Adaptive MPC Path Tracking Control Study of Rear-Wheel-Steering Agricultural Machinery
Reprinted from: *Agriculture* **2024**, *14*, 823, https://doi.org/10.3390/agriculture14060823 . . . . . **49**

Chenbo Shi, Yuanzheng Mo, Xiangqun Ren, Jiahao Nie, Chun Zhang, Jin Yuan and Changsheng Zhu
Improved Real-Time Models for Object Detection and Instance Segmentation for *Agaricus bisporus* Segmentation and Localization System Using RGB-D Panoramic Stitching Images
Reprinted from: *Agriculture* **2024**, *14*, 735, https://doi.org/10.3390/agriculture14050735 . . . . . **70**

Jian Liu, Shihui Yu, Xuemei Liu, Guohang Lu, Zhenbo Xin and Jin Yuan
Super-Resolution Semantic Segmentation of Droplet Deposition Image for Low-Cost Spraying Measurement
Reprinted from: *Agriculture* **2024**, *14*, 106, https://doi.org/10.3390/agriculture14010106 . . . . . **88**

Abdallah E. Elwakeel, Yasser S. A. Mazrou, Ahmed S. Eissa, Abdelaziz M. Okasha, Adel H. Elmetwalli, Abeer H. Makhlouf, et al.
Design and Validation of a Variable-Rate Control Metering Mechanism and Smart Monitoring System for a High-Precision Sugarcane Transplanter
Reprinted from: *Agriculture* **2023**, *13*, 2218, https://doi.org/10.3390/agriculture13122218 . . . . **111**

Jinyang Li, Zhenyu Nie, Yunfei Chen, Deqiang Ge and Meiqing Li
Development of Boom Posture Adjustment and Control System for Wide Spray Boom
Reprinted from: *Agriculture* **2023**, *13*, 2162, https://doi.org/10.3390/agriculture13112162 . . . . **131**

Wang Yang, Junhui Xi, Zhihao Wang, Zhiheng Lu, Xian Zheng, Debang Zhang and Yu Huang
Embedded Field Stalk Detection Algorithm for Digging–Pulling Cassava Harvester Intelligent Clamping and Pulling Device
Reprinted from: *Agriculture* **2023**, *13*, 2144, https://doi.org/10.3390/agriculture13112144 . . . . **161**

Kuo Yang, Xinhui Liu, Changyi Liu and Ziwei Wang
Motion-Control Strategy for a Heavy-Duty Transport Hexapod Robot on Rugged Agricultural Terrains
Reprinted from: *Agriculture* **2023**, *13*, 2131, https://doi.org/10.3390/agriculture13112131 . . . . **181**

Yefeng Sun, Liang Gong, Wei Zhang, Bishu Gao, Yanming Li and Chengliang Liu
Drivable Agricultural Road Region Detection Based on Pixel-Level Segmentation with Contextual Representation Augmentation
Reprinted from: *Agriculture* **2023**, *13*, 1736, https://doi.org/10.3390/agriculture13091736 . . . . **204**

Bohan Fu, Weizhong Sun and Zhao Zhang
Optimal Sand−Paving Parameters Determination of an Innovatively Developed Automatic Maize Seeding Machine
Reprinted from: *Agriculture* **2023**, *13*, 1538, https://doi.org/10.3390/agriculture13081538 . . . . **217**

Wenbo Wei, Maohua Xiao, Weiwei Duan, Hui Wang, Yejun Zhu, Cheng Zhai and Guosheng Geng
Research Progress on Autonomous Operation Technology for Agricultural Equipment in Large Fields
Reprinted from: *Agriculture* **2024**, *14*, 1473, https://doi.org/10.3390/agriculture14091473 . . . . **233**

Rebecca Sarku, Ulfia A. Clemen and Thomas Clemen
The Application of Artificial Intelligence Models for Food Security: A Review
Reprinted from: *Agriculture* **2023**, *13*, 2037, https://doi.org/10.3390/agriculture13102037 . . . . **253**

Wei Liu, Shijie Tian, Qingyu Wang and Huanyu Jiang
Key Technologies of Plug Tray Seedling Transplanters in Protected Agriculture: A Review
Reprinted from: *Agriculture* **2023**, *13*, 1488, https://doi.org/10.3390/agriculture13081488 . . . . **281**

About the Editors

Jin Yuan

Jin Yuan is a professor and doctoral supervisor at the School of Mechanical and Electrical Engineering, Shandong Agricultural University, and a deputy director of the Artificial Intelligence Branch of the Chinese Society of Agricultural Machinery. He is a postdoctoral fellow at Shanghai Jiaotong University and is engaged in scientific cooperation research at the Norwegian University of Science and Technology. He has presided over more than 20 scientific research projects in the fields of agricultural robotics, intelligent agricultural equipment, and precision agriculture, such as the National Natural Science Foundation of China and the National Key Research and Development Program, amongst others. He won two National Science and Technology Advancement Prizes in 2009 and 2011 and was awarded the Third Youth Science and Technology Prize of the Chinese Society of Agricultural Machinery in 2012. He has published more than 60 academic papers and more than 50 of his invention patents have been authorized.

Zichen Huang

Zichen Huang is an assistant researcher at Zhejiang University. He graduated from Kyoto University, Japan, and has been a researcher and guest researcher at Kyoto University. He is mainly engaged in research into intelligent sensing systems and agricultural robots for fruit and vegetable production, including the intelligent sensing of fruit and vegetable growth conditions, the navigation systems of agricultural robots, and multi-factor coupled plant growth modeling. He was selected as a JSPS International Research Fellow in 2021. He has published 13 SCI papers and 3 EI papers as the first or corresponding author. He is currently serving as a youth Editorial Board Member of the *Journal of Integrative Agriculture* and *Smart Agriculture*, and has chaired four sub-forums of IEEE IROS, CIGR, and other international conferences.

Preface

It is with great pleasure that we present this Special Issue, *Intelligent Agricultural Machinery and Robots: Embracing Technological Advancements for a Sustainable and Highly Efficient Agricultural Future*. This collection of fourteen outstanding research papers reflects the latest advancements in intelligent machinery, robotics, and AI-driven technologies in agriculture. In the face of increasing global challenges such as labor shortages, environmental concerns, and rising food demand, these innovative contributions illustrate the pivotal role of technology in shaping the future of sustainable agriculture.

The research presented in this Special Issue spans a broad spectrum of innovative approaches. Autonomous agricultural machinery and robotics are at the forefront, showcasing advancements in path planning, multi-robot collaboration, terrain navigation, and adaptive control systems. These technologies highlight the potential of automation to enhance efficiency, optimize resource use, and reduce environmental impacts across diverse farming environments.

Sensor-based precision agriculture emerges as another key focus, with studies demonstrating the application of machine learning and imaging technologies for tasks such as disease detection, crop monitoring, and irrigation optimization. These solutions underline the promise of data-driven approaches in achieving sustainable crop management practices that conserve resources and minimize chemical inputs.

This Special Issue also emphasizes the increasing integration of robotic systems to automate labor-intensive agricultural tasks, from transplanting seedlings to managing transportation and seeding operations. By streamlining processes and improving consistency, these systems address critical labor shortages while enabling scalability in agricultural production.

Food safety and quality assurance are explored through innovations in monitoring technologies, ensuring product integrity from farm to table. Furthermore, a review of artificial intelligence applications sheds light on their transformative impact on global food security, emphasizing the importance of inclusive and localized approaches to technology adoption.

We extend our gratitude to the authors, reviewers, and editorial staff who contributed to this Special Issue. Together, these studies form a robust foundation for further innovation in intelligent agricultural systems, advancing the goal of a sustainable and highly efficient agricultural future.

Jin Yuan and Zichen Huang
Guest Editors

agriculture

Editorial

Intelligent Agricultural Machinery and Robots: Embracing Technological Advancements for a Sustainable and Highly Efficient Agricultural Future

Jin Yuan [1,*] and Zichen Huang [2,3,*]

[1] School of Mechanical and Electronic Engineering, Shandong Agriculture University, Taian 271018, China
[2] College of Biosystems Engineering and Food Science, Zhejiang University, Hangzhou 310058, China
[3] The Rural Development Academy & Agricultural Experiment Station, Zhejiang University, Hangzhou 310058, China
* Correspondence: jyuan@sdau.edu.cn (J.Y.); huangzichenss@zju.edu.cn (Z.H.)

Agriculture today stands on the brink of transformative innovation, driven by technological advancements in intelligent machinery and robotics. In the context of rising global food demands, labor shortages, and the urgency for sustainable practices, integrating these advanced technologies is no longer a choice but a necessity. This Special Issue, "Intelligent Agricultural Machinery and Robots: Embracing Technological Advancements for a Sustainable and Highly Efficient Agricultural Future", gathers research that underscores the multifaceted roles of robotics, artificial intelligence (AI), and precision sensing in agricultural applications. The assembled papers cover a spectrum of applications, from autonomous field operations and crop management to food security. Together, they demonstrate significant technical strides while laying a solid foundation for modern agriculture's new paradigm.

One prominent theme within this Special Issue is the continued evolution of path-planning and control mechanisms that enhance autonomous agricultural operations. Four papers offer insights into advancements that boost precision and adaptability [1–4]. Efficient path planning for headland turns optimizes field coverage, significantly reducing fuel consumption and minimizing soil compaction [1]. Multi-robot collaboration models support cooperative task allocation and route planning, optimizing resource use and efficiency [3]. Additionally, 3D LiDAR-based crop row detection enables precise navigation within rows, enhancing accuracy in varied terrains [2]. Moreover, a model predictive control approach for rear-wheel steering demonstrates robust adaptability, as it dynamically adjusts parameters for path tracking under different field conditions [4]. Together, these advancements elevate the precision and efficiency of autonomous agricultural machinery, addressing key challenges of modern mechanized farming.

Sensor-based monitoring and precision agriculture technologies play a pivotal role in sustainable crop management, as demonstrated by four studies in this Special Issue [5–8]. Neural networks applied to real-time fruit detection enable effective passion fruit identification and counting [5], showcasing the promise of machine vision for crop yield estimation. Hyperspectral imaging for disease detection [6] demonstrates the power of spectral analysis for the early identification of tea leaf diseases, supporting targeted interventions with reduced chemical inputs. Soil moisture prediction models further illustrate the value of machine learning in optimizing irrigation schedules, conserving water while maximizing productivity [7]. Super-resolution semantic segmentation of spray droplets offers a low-cost solution for precision spraying, ensuring accurate droplet deposition and reducing pesticide use [8]. These advancements support precise crop health monitoring, leading to more efficient and environmentally conscious agricultural practices.

As labor-intensive agricultural tasks increasingly become automated, robotic systems are now integral to the industry's productivity. This theme is explored [9–11], with each

Citation: Yuan, J.; Huang, Z. Intelligent Agricultural Machinery and Robots: Embracing Technological Advancements for a Sustainable and Highly Efficient Agricultural Future. *Agriculture* **2024**, *14*, 2150. https://doi.org/10.3390/agriculture14122150

Received: 18 November 2024
Accepted: 21 November 2024
Published: 26 November 2024

Copyright: © 2024 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (https://creativecommons.org/licenses/by/4.0/).

paper focusing on specialized robots for specific agricultural functions. An adaptive tomato seedling transplanter automates seedling placement [9], minimizing labor needs while ensuring planting consistency. A hexapod transport robot addresses terrain challenges, enabling reliable transport across rugged landscapes [10], which is essential for remote or uneven farmland. The automatic maize seeding machine optimizes the sand-paving process during seeding [11], offering a scalable solution for large-scale farms. These studies illustrate robotics' versatility in addressing various agricultural challenges, showcasing how precision automation can replace repetitive tasks, improve consistency, and make large-scale farming more manageable.

In addition to production-focused technologies, this Special Issue also explores food safety and quality assurance innovations [12,13]. Microcrack detection in eggs, powered by a ConvNext-based U-Net, enhances food quality monitoring along the supply chain, using rapid and accurate detection to prevent product losses [12]. Complementing this, a review of large-scale autonomous equipment outlines the scalability of unmanned operations in meeting the needs of extensive fields, further underscoring the potential of autonomous solutions for efficient, high-quality agricultural management [13]. Together, these studies highlight how precise monitoring and control can ensure product integrity from farms to consumers, reinforcing the role of advanced sensing in maintaining food quality.

Finally, AI's potential in enhancing global food security is explored [14], which delves into the socio-economic impacts of AI-driven solutions across various regions. This review emphasizes the need for iterative, localized approaches to AI-driven food security research, involving stakeholders to ensure models' real-world validity and sustainability. This holistic approach underscores AI not merely as an optimization tool but as a transformative technology capable of supporting resilient food systems worldwide, particularly in regions facing socio-economic and environmental challenges.

The collection of studies reflects a convergence of robotics, AI, and sensor technology aimed at elevating agricultural efficiency, precision, and sustainability. The advancements showcased here, from intelligent control of autonomous vehicles to deep learning applications in crop monitoring, signify a shift toward a future where intelligent agricultural machinery and robots play an indispensable role. By embracing these innovations, agriculture moves closer to a sustainable, high-efficiency future that aligns with environmental and societal imperatives, ultimately building a foundation for a resilient global food system.

Author Contributions: Conceptualization, J.Y.; investigation, Z.H.; writing—original draft preparation, Z.H.; writing—review and editing, J.Y. All authors have read and agreed to the published version of the manuscript.

Funding: This material is based upon work that is supported by the National Natural Science Foundation of China (Grant No. 52275262).

Institutional Review Board Statement: Not applicable.

Data Availability Statement: Not applicable.

Acknowledgments: We would like to sincerely thank all the authors who submitted papers to this Special Issue of *Agriculture* entitled "Intelligent Agricultural Machinery and Robots", the reviewers of these papers for their constructive comments and thoughtful suggestions, and the editorial staff of *Agriculture*.

Conflicts of Interest: The authors declare no conflicts of interest.

References

1. Yang, S.; Zhang, J.; Yuan, J. A High-Accuracy Contour Segmentation and Reconstruction of a Dense Cluster of Mushrooms Based on Improved SOLOv2. *Agriculture* **2024**, *14*, 1646. [CrossRef]
2. Li, J.; Nie, Z.; Chen, Y.; Ge, D.; Li, M. Development of Boom Posture Adjustment and Control System for Wide Spray Boom. *Agriculture* **2023**, *13*, 2162. [CrossRef]

3. Elwakeel, A.E.; Mazrou, Y.S.A.; Eissa, A.S.; Okasha, A.M.; Elmetwalli, A.H.; Makhlouf, A.H.; Metwally, K.A.; Mahmoud, W.A.; Elsayed, S. Design and Validation of a Variable-Rate Control Metering Mechanism and Smart Monitoring System for a High-Precision Sugarcane Transplanter. *Agriculture* **2023**, *13*, 2218. [CrossRef]
4. Wang, M.; Niu, C.; Wang, Z.; Jiang, Y.; Jian, J.; Tang, X. Model and Parameter Adaptive MPC Path Tracking Control Study of Rear-Wheel-Steering Agricultural Machinery. *Agriculture* **2024**, *14*, 823. [CrossRef]
5. Sun, Y.; Gong, L.; Zhang, W.; Gao, B.; Li, Y.; Liu, C. Drivable Agricultural Road Region Detection Based on Pixel-Level Segmentation with Contextual Representation Augmentation. *Agriculture* **2023**, *13*, 1736. [CrossRef]
6. Yang, W.; Xi, J.; Wang, Z.; Lu, Z.; Zheng, X.; Zhang, D.; Huang, Y. Embedded Field Stalk Detection Algorithm for Digging–Pulling Cassava Harvester Intelligent Clamping and Pulling Device. *Agriculture* **2023**, *13*, 2144. [CrossRef]
7. Liu, W.; Tian, S.; Wang, Q.; Jiang, H. Key Technologies of Plug Tray Seedling Transplanters in Protected Agriculture: A Review. *Agriculture* **2023**, *13*, 1488. [CrossRef]
8. Liu, J.; Yu, S.; Liu, X.; Lu, G.; Xin, Z.; Yuan, J. Super-Resolution Semantic Segmentation of Droplet Deposition Image for Low-Cost Spraying Measurement. *Agriculture* **2024**, *14*, 106. [CrossRef]
9. Shi, C.; Mo, Y.; Ren, X.; Nie, J.; Zhang, C.; Yuan, J.; Zhu, C. Improved Real-Time Models for Object Detection and Instance Segmentation for Agaricus Bisporus Segmentation and Localization System Using RGB-D Panoramic Stitching Images. *Agriculture* **2024**, *14*, 735. [CrossRef]
10. Yang, K.; Liu, X.; Liu, C.; Wang, Z. Motion-Control Strategy for a Heavy-Duty Transport Hexapod Robot on Rugged Agricultural Terrains. *Agriculture* **2023**, *13*, 2131. [CrossRef]
11. Fu, B.; Sun, W.; Zhang, Z. Optimal Sand−Paving Parameters Determination of an Innovatively Developed Automatic Maize Seeding Machine. *Agriculture* **2023**, *13*, 1538. [CrossRef]
12. Shi, C.; Li, Y.; Jiang, X.; Sun, W.; Zhu, C.; Mo, Y.; Yan, S.; Zhang, C. Real-Time ConvNext-Based U-Net with Feature Infusion for Egg Microcrack Detection. *Agriculture* **2024**, *14*, 1655. [CrossRef]
13. Wei, W.; Xiao, M.; Duan, W.; Wang, H.; Zhu, Y.; Zhai, C.; Geng, G. Research Progress on Autonomous Operation Technology for Agricultural Equipment in Large Fields. *Agriculture* **2024**, *14*, 1473. [CrossRef]
14. Sarku, R.; Clemen, U.A.; Clemen, T. The Application of Artificial Intelligence Models for Food Security: A Review. *Agriculture* **2023**, *13*, 2037. [CrossRef]

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.

agriculture

MDPI

Article

Real-Time ConvNext-Based U-Net with Feature Infusion for Egg Microcrack Detection

Chenbo Shi, Yuejia Li, Xin Jiang, Wenxin Sun, Changsheng Zhu, Yuanzheng Mo, Shaojia Yan and Chun Zhang *

College of Intelligent Equipment, Shandong University of Science and Technology, Tai'an 271019, China; skd996523@sdust.edu.cn (C.S.); 202283230037@sdust.edu.cn (Y.L.); jx1093560806@163.com (X.J.); sunwenx2001@163.com (W.S.); cs.zhu@sdust.edu.cn (C.Z.); 202283230023@sdust.edu.cn (Y.M.); 202383230010@sdust.edu.cn (S.Y.)

* Correspondence: zhangchun@sdust.edu.cn

Abstract: Real-time automatic detection of microcracks in eggs is crucial for ensuring egg quality and safety, yet rapid detection of micron-scale cracks remains challenging. This study introduces a real-time ConvNext-Based U-Net model with Feature Infusion (CBU-FI Net) for egg microcrack detection. Leveraging edge features and spatial continuity of cracks, we incorporate an edge feature infusion module in the encoder and design a multi-scale feature aggregation strategy in the decoder to enhance the extraction of both local details and global semantic information. By introducing large convolution kernels and depth-wise separable convolution from ConvNext, the model significantly reduces network parameters compared to the original U-Net. Additionally, a composite loss function is devised to address class imbalance issues. Experimental results on a dataset comprising over 3400 graded egg microcrack image patches demonstrate that CBU-FI Net achieves a reduction in parameters to one-third the amount in the original U-Net, with an inference speed of 21 ms per image (1 million pixels). The model achieves a Crack-IoU of 65.51% for microcracks smaller than 20 μm and a Crack-IoU and MIoU of 60.76% and 80.22%, respectively, for even smaller cracks (less than 5 μm), achieving high-precision, real-time detection of egg microcracks. Furthermore, on the publicly benchmarked CrackSeg9k dataset, CBU-FI Net achieves an inference speed of 4 ms for 400 × 400 resolution images, with an MIoU of 81.38%, proving the proposed method's robustness and generalization capability across various cracks and complex backgrounds.

Keywords: poultry eggs; microcrack detection; lightweight model; semantic segmentation; deep learning

Citation: Shi, C.; Li, Y.; Jiang, X.; Sun, W.; Zhu, C.; Mo, Y.; Yan, S.; Zhang, C. Real-Time ConvNext-Based U-Net with Feature Infusion for Egg Microcrack Detection. *Agriculture* **2024**, *14*, 1655. https://doi.org/10.3390/agriculture14091655

Academic Editor: Massimo Cecchini

Received: 27 July 2024
Revised: 16 September 2024
Accepted: 19 September 2024
Published: 22 September 2024

Copyright: © 2024 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (https://creativecommons.org/licenses/by/4.0/).

1. Introduction

In 2023, China's poultry egg production and consumption reached approximately 34.5 million tons. Poultry eggs are rich in nutrients and have a significant positive impact on human health and development. The poultry egg industry plays a crucial role in driving national economic growth, promoting agricultural development, and ensuring food safety. However, during production, storage, and distribution, poultry eggs are prone to defects such as cracks, damage, and contamination [1], which typically require manual inspection and sorting. Microcracks, in particular, severely affect egg quality but are often overlooked during inspection [2]. Eggshell damage is a critical issue in egg production and a major source of economic loss for the industry [3]. As a result, researchers both domestically and internationally are increasingly focused on developing high-speed, high-precision poultry egg crack detection technologies. Current research primarily centers on acoustic, optical, and electrical analysis, as well as deep learning methods, achieving notable advancements in these fields.

The acoustic-based eggshell crack detection method leverages the physical properties of sound to analyze the vibration signals generated by striking an eggshell, offering a

novel research approach. Eggshell cracks can affect the structural strength and damping coefficient of eggs, which are reflected in the frequency and intensity of vibration signals. Chia-Chun Lai et al. [4] used cross-correlation analysis and Bayesian classification to classify and detect acoustic response signals generated by eggshell cracks, achieving a detection rate of 97%. Li Sun et al. [5] employed acoustic resonance analysis to detect eggshell cracks by impacting eggshells and collecting their response signals. They calculated Pearson correlation coefficients (PCCs) as feature parameters to compare the signal similarity between intact and cracked eggs at various sampling points. Based on these feature parameters, they established a general linear discriminant function to distinguish between intact and cracked eggs. In their experiment, a mixed sample of 100 chicken eggs and 100 duck eggs was used, and the crack detection rate reached 95.5%. However, these methods are susceptible to environmental conditions, egg shapes, crack locations and impact angles, leading to insufficient detection stability. Moreover, excessive impact force may further damage the eggshell structure. Additionally, since microcracks smaller than 5 mm have minimal impacts on the structural rigidity and resonance characteristics of eggs, acoustic techniques achieve lower accuracy in detecting microcracks [6].

Shi Chenbo et al. [7] proposed a non-destructive detection method for eggshell cracks based on a model of the electrical characteristics of eggs. This method involves establishing an electric field on the eggshell surface and analyzing the subtle current changes caused by cracks in the eggshell. The detection hardware system is designed based on electrical characteristic analysis. Additionally, Shi Chenbo et al. [8] developed a wavelet-scattering convolutional network algorithm utilizing 1D-CNN, LSTM, BiLSTM, and GRU networks for the non-destructive detection of microcracks in eggshells from electrical signals. While this method achieves high accuracy, it has limitations in detection coverage, failing to comprehensively identify all microcracks. Additionally, air humidity can affect conductivity, introducing instability to this method. The detection performance of this method may also be suboptimal for hidden cracks that have not fully penetrated the eggshell or for dry, fine cracks.

Compared to acoustic and electrically based methods, optically based egg crack detection methods offer the advantages of being non-destructive and stable. This technique captures images of the eggshell surface using image acquisition devices and processes and analyzes the image data. Therefore, machine vision-based eggshell crack detection methods are widely used in practical applications and have developed relatively mature technologies [9]. Guanjun Bao et al. [10] addressed irregular dark spots and invisible microcracks on the eggshell surface by using a negative LOG (Laplacian of Gaussian) operator to enhance cracks. They then employed a hysteresis threshold algorithm to eliminate irrelevant dark spots in the binary image, ensuring crack continuity. Finally, they used an improved Local Fitting Image (LFI) indicator to distinguish between cracks and false markings, achieving a crack detection rate of 92.5%. Kunshan Yao et al. [11] proposed methods such as crack enhancement and dual-threshold segmentation to extract the geometric features of cracks. Using the XGBoost classification model, they classified cracked eggs and achieved a recognition accuracy of 93.33%.

Traditional machine vision methods can effectively detect eggshell cracks; however, they often suffer from algorithmic complexity and slow processing speeds. With the recent advancements in deep learning in the field of computer vision, deep learning-based approaches have demonstrated superior performance in crack detection. Currently, the primary research achievements in eggshell crack detection focus on image classification tasks. Bhavya Botta et al. [12] created an eggshell crack dataset containing 468 images, with crack widths ranging from 20 to 40 μm and an average length of 30 mm. By employing convolutional neural networks (CNNs), they achieved a detection accuracy of 95.38%. Wenquan Tang [13] proposed the MobileNetV3_egg model, based on MobileNetV3_large, for real-time detection of damaged preserved eggs, achieving an accuracy of 96.3% and detecting 300 images in 4.267 s. Amin Nasiri et al. [14] collected images of intact, broken, and blood-stained eggs and used a CNN classification algorithm based on the VGG16

architecture, achieving an accuracy of 94.85%. With the introduction of segmentation networks such as U-Net, DeepLab-V3, and PSPNet, improved semantic segmentation models have also been applied to detection tasks. Xiuying Xu et al. [15] developed a semantic segmentation model based on an improved U-Net for the segmentation of corn stalks in fields with complex backgrounds, achieving an accuracy of 93.87%, outperforming U-Net, SegNet, and ResNet models. Chengqi Liu et al. [16] proposed a DeepLab V3+-based deep learning method for pig image segmentation in small sample sizes. They optimized the fusion of high- and low-frequency features using a recursive cascading approach to extract latent semantic information, resulting in a single-label model MIoU of 76.31%. However, these models are complex, with a large number of parameters and high training costs; therefore, they may not be suitable for the detection of narrow microcracks with few pixels in the images.

In the context of microcrack detection in poultry eggs for industrial applications, it is essential to address the model's scale and efficiency. Given the constraints of computational power, a pressing research issue is how to reduce the parameters of semantic segmentation models to make them more lightweight and accelerate inference without compromising detection accuracy. In 2022, Zhuang Liu et al. [17] redesigned a purely convolutional neural network model based on the standard ResNet and named it ConvNeXt. This model outperformed the Swin Transformer in the field of computer vision, with lower computational costs. Zhimeng Han et al. [18] proposed an effective model combining U-Net and ConvNeXt for medical image segmentation, achieving leading results with fewer parameters. Inspired by ConvNeXt, our model introduces improvements such as large convolution kernels and depthwise separable convolutions to reduce the number of model parameters.

In microcrack detection, challenges arise due to the narrow width of cracks and the similarity in contrast between the crack ends and the background texture. This paper addresses these challenges through the following two approaches: by enhancing detailed features and fully aggregating feature information for segmentation. In segmentation tasks, cracks can be divided into edge and internal regions [19]. In the field of object detection, some researchers [20] have demonstrated that utilizing and integrating edge features can improve the model's perception accuracy. Crack edge features help the model distinguish and locate crack defects. In digital image processing, edge detection effectively extracts object edge information by identifying pixels with significant brightness changes. Common detection algorithms include the Sobel, Prewitt, and Canny algorithms. Rasha Alshawi et al. [21] proposed using Sobel and Canny filters for edge detection to extract engineering features and manually integrate them into the network model to guide segmentation. This feature enhancement method provides a robust solution for segmentation tasks on limited and skewed datasets. Inspired by data augmentation, this paper's model incorporates a strategy of injecting image edge features into the encoder design to expand the detail feature space, thereby enhancing the model's ability to detect microcracks.

Additionally, Hengshuang Zhao et al. [22] proposed the Pyramid Scene Parsing Network (PSPNet) for pixel-level prediction tasks, providing an excellent framework by aggregating context from different regions. Building on this, Guosheng Lin et al. [23] proposed a multi-stage refinement network (RefineNet) that effectively integrates missing information from downsampling, producing high-resolution predictive images. It is evident that multi-scale feature fusion is beneficial for improving segmentation accuracy. Therefore, our model aggregates multi-scale features in the decoder to obtain fused features. Ablation experiments demonstrate that this approach significantly enhances accuracy compared to single upsampling of high-resolution feature maps for prediction.

In microcrack image segmentation, addressing class imbalance is a critical challenge due to the uneven distribution of foreground and background in the images. The positive–negative cross-entropy loss function is commonly used for most semantic segmentation tasks. However, when foreground pixels are significantly fewer than background

pixels, the background elements dominate the loss function calculation. This dominance causes the model to become overly biased toward the background, adversely affecting training and prediction outcomes. The Dice coefficient-based loss function proposed by Fausto Milletari et al. [24] has been widely adopted to mitigate the negative impact of imbalanced foreground and background areas on model performance. The Dice loss function places more emphasis on capturing the foreground regions, ensuring a lower false-negative rate. However, it suffers from the issue of loss saturation and is typically used in conjunction with the cross-entropy loss function. Michael Yeung et al. [25] proposed a unified focal loss that combines cross-entropy loss and Dice loss to address class imbalance. Experimental results indicate that this loss function exhibits robust performance in handling class imbalance.

In summary, this paper proposes a lightweight and highly real-time algorithm for the detection of microcracks in poultry eggs. The main innovations of this study are summarized as follows:

- This paper proposes a Real-time ConvNext-Based U-Net architecture with Feature Infusion for egg microcrack detection (CBU-FI Net). This architecture leverages the strengths of both U-Net and ConvNeXt, using ConvNeXt's fundamental modules as the backbone network. This approach significantly reduces the model's parameter count and computational complexity.
- To address the challenges in microcrack detection tasks, this paper introduces a feature infusion module within the encoder and employs multi-scale feature aggregation in the decoder for segmentation. By incorporating edge information, this strategy expands the spatial representation of crack features and enhances the extraction of both local details and global semantic information. This approach significantly improves microcrack segmentation accuracy, even with limited training data.
- To tackle the challenge of positive and negative sample imbalance in microcrack images for practical industrial applications, this paper introduces a hybrid loss function that combines cross-entropy loss and Dice loss. This approach significantly improves segmentation performance on microcrack images.

2. Materials and Methods

2.1. Detection Framework

This paper proposes CBU-FI Net, a real-time U-Net architecture enhanced with feature infusion, to meet the requirements for accurate detection and real-time performance in online poultry egg inspection. Figure 1 shows the overall network framework. With a parameter count of approximately 28 million, U-Net is considered a relatively lightweight segmentation model. In this study, U-Net is selected as the baseline model, then improved upon. Given that microcracks are characterized by small detection targets and low contrast against background textures, the primary focus of these improvements is on enhancing the network's ability to detect limited microcrack information and achieving efficient and rapid segmentation tasks with minimal parameters and computational load.

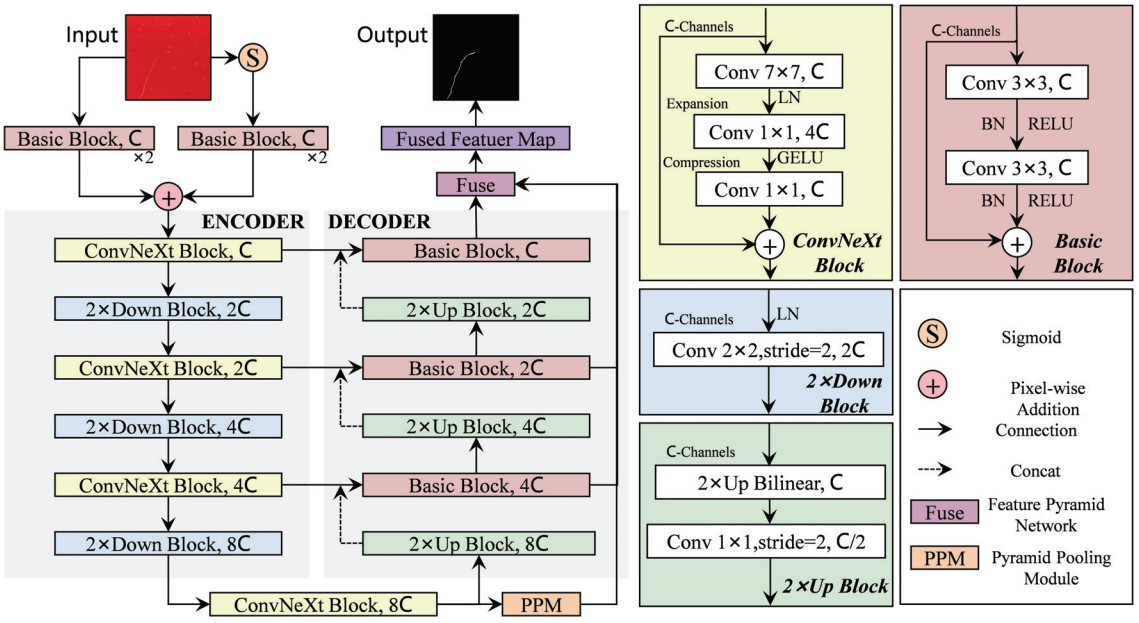


Figure 1. Architectural design of CBU-FI Net. The network is composed of three levels of encoders and decoders. The encoder includes three ConvNeXt blocks and downsampling layers, which progressively capture multi-scale features. The decoder consists of three basic blocks and upsampling layers that restore spatial resolution and integrate multi-scale features, enabling precise microcrack segmentation.

2.1.1. Overall Network Framework Design

Building upon the original U-Net architecture, CBU-FI Net achieves high-precision and high-speed semantic segmentation through its encoder–decoder framework. The improvement strategies for the encoder in this model are outlined as follows:

1. Drawing inspiration from dual-branch models, the traditional U-Net architecture is augmented with a feature infusion branch, enhancing the feature-space information of crack edges. To minimize the computational load of the fully convolutional network, crack images are processed through a dual-branch path before entering the encoder. Each branch employs a basic block to extract features from both the original image and the edge image. Consequently, the number of image channels increases from 3 to 48, and the image dimensions are reduced to one-fourth of their original size. The outputs from the two branches are then combined through pixel-wise summation before being input into the encoder.
2. The original U-Net architecture involves four downsampling operations, resulting in 1024 channels. To reduce model parameters and computational complexity, this model differs from the original U-Net by using ConvNeXt basic blocks as the backbone, replacing the traditional CBR structure with convolution operations. Additionally, the number of downsampling operations is reduced to three, and the number of channels is decreased to 384.
3. The original U-Net architecture employs max pooling for downsampling, which results in information loss—a significant drawback for microcrack segmentation tasks. To address this issue, the proposed model utilizes learnable dilated convolutions with a kernel size of 2 and a stride of 2 for downsampling. After three downsampling operations, the resolution of the feature maps is reduced to 1/32 of the original image size.

In the decoder stage, basic blocks from ResNet replace the two 3×3 convolutional layers of the original U-Net architecture, and bilinear interpolation is used for upsampling. In the final layer, a PPM (Pyramid Pooling Module) is applied to aggregate global context, expanding the receptive field, while an FPN (Feature Pyramid Network) combines multi-scale feature maps from basic blocks at various levels of the network. This approach allows the model to effectively capture both global and local context, which is crucial for accurately detecting and segmenting cracks in eggs. The multi-scale fusion of features, compared to using only the single highest-resolution feature map, minimizes information loss and enhances the precision of segmentation, particularly for small cracks and intricate boundary details.

2.1.2. Lightweight Trunk Feature Extraction Networks

ConvNeXt is a hybrid model that improves upon previous modules, leading to CNN performance in computer vision that surpasses that of the Swin Transformer. Specifically, it limits computational cost. In this model, the original CBR (Conv + BatchNorm + ReLU) structure is replaced with a more lightweight ConvNeXt Block, enhancing training and inference speed. As illustrated, the ConvNeXt Block begins with a 7×7 depthwise separable convolution, which captures more crack detail than a 3×3 convolution. Next, the expansion layer uses a 1×1 convolution to increase feature dimensions, followed by a compression layer, which also employs a 1×1 convolution to reduce the number of channels in the feature map, thereby decreasing the model's parameters. The output of the ConvNeXt block is a multi-channel feature map derived from the feature map output by the previous convolution module or its downsampled version. The feature maps output at various scales during the U-Net backbone stage are substituted with those produced by the ConvNeXt block at corresponding scales. Consequently, in the skip connection stage, these feature maps can still be concatenated with the upsampled feature maps, which are rich in abstract semantic information.

2.1.3. Feature Infusion Module

In microcrack detection, challenges arise due to the narrow width of the cracks and the low contrast between crack tips and the background texture. Edge features of cracks assist the model in distinguishing and locating crack defects. Therefore, incorporating edge information helps achieve clearer and more accurate crack shapes. Inspired by these methods, this paper introduces a feature infusion module based on the intrinsic characteristics of cracks. This module is integrated into the deep learning model to introduce edge information into crack detection, expanding the feature-space information of cracks. Consequently, it provides additional guidance to enhance model performance, particularly in cases of data imbalance.

The model utilizes the computationally efficient Sobel operator as an edge detector to achieve real-time crack detection, meeting the requirements for real-time processing. Specifically, prior to the encoder, two 3×3 convolution kernels are employed to implement the Sobel edge detector, which detects crack edges and highlights crack contours to generate an edge map of the crack image. To align the features of the original image with those of the edge image, a dual-branch fusion strategy is adopted. This strategy uses identical basic feature extraction blocks to align and extract features from both the original and edge images. The features from these dual branches are then combined using element-wise addition, effectively integrating the extracted edge features into the segmentation process.

Seg-Grad-CAM is employed to visualize class activation maps [26] to better understand the impact of feature guidance on crack segmentation. Under identical conditions, feature heat maps of models with and without the feature infusion structure are compared on a microcrack dataset. With feature infusion, the model exhibits increased focus on fine cracks, particularly enhancing the prediction accuracy for crack tips. In contrast, without feature infusion, the model struggles to detect small crack tips. This illustrates that the dual-branch feature infusion strategy significantly expands the feature space, surpassing

the segmentation capabilities of conventional deep learning and mitigating the limitations posed by class imbalance (Figure 2).

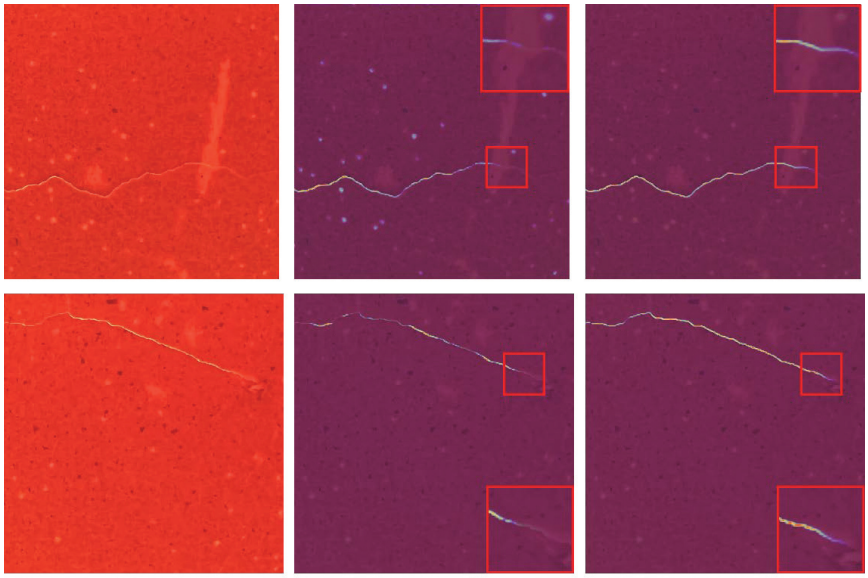


Figure 2. A visual comparison of the models with and without the feature infusion structure. The smaller red boxes in the figure highlights the end of the crack, and the zoomed-in view is displayed in the top-right/lower-right corner of the figure. By magnifying the segmentation results of microcrack tips, the figure emphasizes the significant advantages of the feature infusion structure in enhancing the accuracy of microcrack detection.

2.1.4. Composite Loss Function

In the training dataset, the imbalance between positive and negative samples is a significant issue. Table 1 shows that in eggshell crack images, the proportion of foreground (crack) pixels relative to background pixels is uneven, with the number of background pixels being 117 times that of crack pixels. This imbalance can cause deep learning models to overly focus on background pixels during training, leading to fluctuations in the loss function, affecting gradient stability, and hindering the improvement of model accuracy. Therefore, measures must be taken to address the imbalance between positive and negative samples to enhance model performance and accuracy.

Table 1. Average percentage of foreground and background pixels relative to total pixels in a cracked image.

Classes	Background	Crack
Percentage value (%)	99.441	0.559

The cross-entropy loss function is one of the most commonly used classification loss functions, measuring the discrepancy between the predicted values and the true labels. The definition of cross-entropy loss is shown in Equation (1).

$$L_{CE} = -\sum^{i} y_i \log \hat{y}_i, \quad (1)$$

where y_i represents the true label; $\hat{y}_i$ represents the probability value predicted by the model, which ranges between 0 and 1; and the log function is the natural logarithm.

In this study, the number of microcrack pixels (target pixels) to be detected is significantly smaller than the number of background pixels, leading to a class imbalance problem. Dice loss, a region-related loss function, performs well in scenarios with class imbalance. Dice loss is defined by Equation (2), denoted as 1-Dice_coefficient, and represents the similarity between two samples.

$$L_{Dice} = 1 - Dice_{coefficient} = 1 - \frac{2 \times |X \cap Y|}{|X| + |Y|}, \tag{2}$$

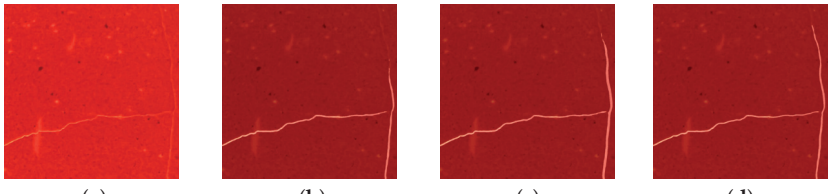
It is computed by the intersection of the background pixel set (X) and the target pixel set (Y). A higher Dice coefficient indicates a more significant similarity between the two samples.

The network's performance in detecting microcracks is improved by proposing a composite loss function that integrates both cross-entropy loss and Dice loss. This approach leverages the strengths of both loss functions, addressing precision issues related to the small proportion of crack pixels and enhancing predictive accuracy. The expression for this composite loss function is provided in Equation (3).

$$Loss = L_{CE} + \alpha L_{Dice}, \tag{3}$$

where α denotes the adjustment factor.

Analysis of the segmentation results in Figure 3b shows that the original image in Figure 3a contains both large cracks and fine microcracks. When relying solely on cross-entropy loss, the model primarily focuses on segmenting prominent main cracks while neglecting smaller or less noticeable microcracks. Thus, cross-entropy loss mainly benefits the model in learning features related to background pixels but performs relatively poorly in capturing fine details. In contrast, as shown in Figure 3c, the model trained with Dice loss exhibits improved segmentation of details and local features, showing higher sensitivity to boundary details and better capturing the complex information related to microcracks. Finally, the composite loss function used in Figure 3d combines the strengths of both loss functions, resulting in superior overall segmentation performance.



(**a**) (**b**) (**c**) (**d**)

Figure 3. Crack segmentation results with different loss functions: (**a**) original image; (**b**) cross-entropy loss alone; (**c**) Dice loss alone; (**d**) composite loss function.

2.2. Dataset

2.2.1. Microcrack Dataset for Eggs

In previous work, we created a large-scale pixel-level annotated dataset comprising 3436 samples of microcrack defects in eggs, with each image having a size of 1024 × 1024 pixels. The crack types in these image blocks include curved, branched, and reticulated uneven cracks. Microcracks predominantly appear at the terminals and branch ends of the cracks, and they are mainly distributed along the edges of the feature maps. Example images from the dataset are shown in Figure 4. In this study, we quantitatively classified the dataset based on the width and average length of the cracks, dividing it into two levels—L1 patch and L2 patch—with detailed information provided in Table 2.

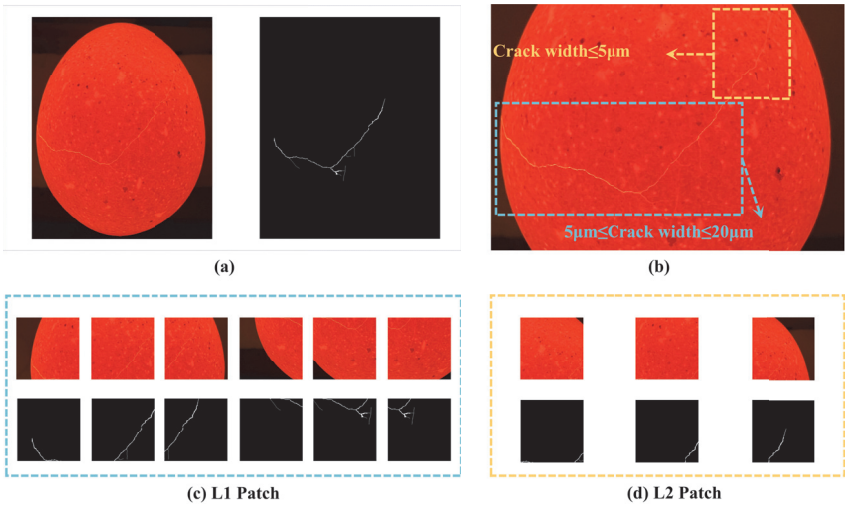


Figure 4. Example images from the dataset: (**a**) original egg image and corresponding mask map; (**b**) crack grading map; (**c**) L1 patch (crack width less than 20 μm and more than 5 μm); (**d**) L2 patch (crack width less than 5 μm).

Table 2. Crack size grading quantization table.

Level	Crack Width	Average Crack Length	Count
L1 Patch	5 μm∼20 μm	10 mm	2486
L2 Patch	≤5 μm	30 μm	950

2.2.2. Public Segmentation Datasets

We selected the publicly available CrackSeg9k [27] dataset for performance assessment to evaluate the reliability and generalizability of the model. This dataset comprises a total of 9255 images (400 × 400 resolution), combining seven different smaller, open-source datasets. It includes various structural cracks in masonry walls, ceramics, concrete bridge decks, walls, and pavement, featuring microcracks with widths ranging from 0.06 mm to 25 mm. Shreyas Kulkarni et al. [27] optimized the dataset by standardizing the masks across multiple classes, creating CrackSeg9k, the largest, most diverse, and most consistent crack segmentation dataset constructed to date.

2.3. Evaluation Metrics

The performance of the trained model in detecting microcracks in poultry eggs is assessed from the following two perspectives: model performance and real-time efficiency. Model performance is evaluated using metrics such as the confusion matrix, pixel accuracy, crack Intersection over Union (Crack-IoU), and Mean Intersection over Union (MIoU) to gauge classification accuracy and segmentation precision. As shown in Equations (4)–(6), these indices take values in the interval of [0, 1], with higher values indicating better segmentation results. Real-time efficiency is evaluated based on inference time, parameter count, and FLOPs (floating-point operations per second). Higher values in these three metrics indicate greater processing power and lower computational complexity for the model. A comprehensive assessment of these metrics offers a thorough evaluation of the model's effectiveness and efficiency in practical applications.

$$Pixel - Accuracy = \frac{TP + TN}{TP + TN + FP + FN}, \quad (4)$$

$$Crack - IoU = \frac{TP}{TP + FP + FN}, \tag{5}$$

$$MIoU = \frac{1}{N} \sum_{k=1}^{n} IoU_k, \tag{6}$$

where k is the number of categories, IoU_k represents the IoU value for type k, and n represents the total number of categories in the dataset.

TP denotes the number of pixels correctly classified as cracks, FP denotes the number of pixels incorrectly classified as cracks, TN denotes the number of pixels correctly classified as background, and FN denotes the number of pixels incorrectly classified as background.

2.4. Experimental Environment and Parameter Settings

Information about the server hardware and software used in this study is shown in Table 3.

Table 3. Server hardware and software information.

Attribute Name	Attribute Content
CPU	12th Gen Intel(R) Core(TM) i7-12700K
GPU	GeForce RTX 3090
Graphics Memory	48G
Operating System	Ubuntu 22.04.1
Deep Learning Framework	MMSegmentation
MMSegmentation Version	1.2.2
Python Version	3.8
Pytorch Version	1.10.1

The egg microcrack dataset contains a total of 3436 sheets, and the ratio of the division of the training set, the validation set, and the test set is 6:1:3. We employed a learning rate strategy described by the following Equation (7):

$$l_r = b_{lr} \times \left(1 - \frac{e}{\max e}\right)^p \tag{7}$$

where b_{lr} represents the base learning rate, e represents the current iteration number, max e denotes the maximum number of iterations for the model, and p is a hyperparameter used to control the learning rate. The hyperparameter values used for network training are presented in Table 4.

Table 4. Hyperparameter information.

Hyperparameter	Values
l_r	Type = Poly *, p = 0.9
Optimizer	Type = SGD **, b_{lr} = 0.01, Momentum *** = 0.9
Batch size	8
Max Iterations	80,000

* An exponential transformation strategy. ** Stochastic gradient descent. *** The momentum algorithm is employed as the optimization algorithm.

3. Results

3.1. Data Analysis

This paper compares six advanced deep learning-based segmentation architectures to evaluate the performance of the CBU-FI Net model in microcrack detection tasks. Extensive experiments were conducted on a microcrack defect detection dataset for eggs and the publicly available CrackSeg9k segmentation dataset. The selected models include three classic semantic segmentation models (U-Net [28], PSPNet (Res50) [22], and DeepLabV3

(Res50) [29]), as well as three state-of-the-art dual-branch models (BiseNetV2 [30], DDR-Net [31], and HrSegNet-B48 [32]). These comparative experiments aim to comprehensively assess and verify the superior performance of CBU-FI Net in detecting microcracks.

3.1.1. Comparative Evaluation of Model Performance on Ungraded Datasets

In the comparative experiments for egg microcrack detection, CBU-FI Net and six other models were trained and tested on the egg microcrack defect detection dataset with identical proportions, with each image having a resolution of 1024 × 1024. Additionally, this model was compared with our previous research result obtained using the MobileUNet-CBAM model [33]. The experiment evaluated the performance of each model in detecting egg microcracks with a crack width of less than 20 μm, with the results presented in Table 5.

Table 5. Results of different models on the ungraded dataset.

Network Structure	Pixel Accuracy (%)	Crack-IoU (%)	MIoU (%)	Inference Time (s)	Parameters (M)	FLOPs (GFLOPs)
U-Net	71.38	61.56	80.28	0.079	29.06	811.64
PSPNet (Res50)	75.93	64.22	81.84	0.069	48.98	713.76
DeepLabV3 (Res50)	76.30	64.30	81.88	0.091	68.01	1078.57
BiseNetV2	76.84	64.43	81.92	0.007	3.341	49.142
DDRNet	77.31	64.95	82.18	0.009	20.299	71.759
HrSegNet-B48	75.15	64.45	81.94	0.008	5.465	32.672
MobileUNet-CBAM	77.04	65.01	82.24	0.080	14.09	96.38
CBU-FI Net	78.70	65.51	82.47	0.021	7.696	102.88

According to the results on the ungraded dataset, the proposed CBU-FI Net achieved the highest pixel accuracy (78.70%), Crack-IoU (65.51%), and MIoU (82.47%) with a small parameter count (7.696M). Classic semantic segmentation models like U-Net and PSPNet (Res50) underperformed, primarily due to their reliance on standard convolutions, which limit their ability to capture the fine-grained contextual details needed for small crack detection. DeepLab-V3 (Res50) performed better due to the use of atrous convolutions that expanded the receptive field, but this came at the cost of a longer inference time and increased computational complexity, making it inefficient for real-time tasks.

Recent advancements have introduced dual-branch architectures, such as BiseNetV2, DDRNet, and HrSegNet-B48, which aim to balance semantic information and fine detail extraction while reducing computational load. While these networks significantly reduced model size and complexity, their light weights designs led to less accurate segmentation of microcracks. BiseNetV2 and DDRNet achieved fast inference by simplifying feature extraction, but this compromised their ability to detect microcracks under 20 μm. HrSegNet-B48, despite improving detail processing with a more complex encoder and modules, still struggled to achieve the precision needed for microcrack detection when compared to CBU-FI Net.

CBU-FI Net's lightweight design surpassed MobileUNet-CBAM, achieving a scale similar to that of dual-branch networks while offering competitive inference speed and delivering superior segmentation accuracy. Its ability to aggregate multi-scale features allowed it to excel in detecting microcracks under 20 μm, striking a balance between model complexity and segmentation performance.

3.1.2. Comparative Analysis of Models on the Graded Dataset (L2 Patch)

The model's performance was evaluated using the L2 patch dataset, where all crack widths are below 5 μm. The results are presented in Figure 5. The decrease in accuracy compared to Experiment 3.1.1 can be attributed to the dataset's finer microcracks, which cause a loss of detail and texture features in low-resolution or downsampled images. This makes it more challenging for models to capture critical microcrack features.

U-Net and PSPNet struggled due to their inability to effectively detect such small details. While DeepLab-V3 (Res50) benefited from atrous convolutions to capture more global information, it still lacked the precision needed for fine cracks under 5 μm. Lightweight models like BiseNetV2, DDRNet, and HrSegNet-B48 prioritized efficiency but sacrificed fine-grained feature extraction, leading to less accurate segmentation in this context. Despite these challenges, the improved CBU-FI Net showed superior performance by effectively retaining fine textures through multi-scale feature aggregation and enhanced upsampling. This allowed it to consistently detect microcracks with greater accuracy, demonstrating its adaptability and generalization capabilities, even for cracks smaller than 5 μm.

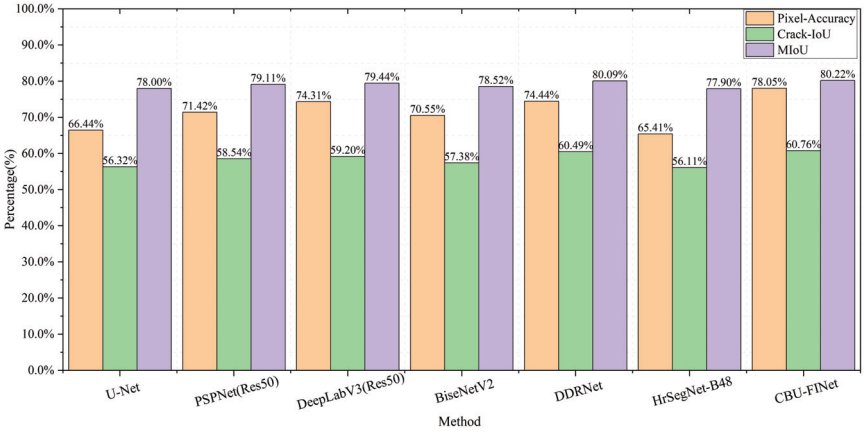


Figure 5. Results of model performance on the L2 patch dataset.

3.1.3. Comparative Analysis of Model Performance on a Public Dataset (CrackSeg9k)

We conducted experiments on the publicly available CrackSeg9k dataset to assess the robustness and generalization capabilities of various semantic segmentation models for crack detection. Our study included CBU-FI Net and six other models. All models were trained and tested on the same proportion of training and testing sets, with each image having a resolution of 400 × 400 pixels. To ensure a valid comparison, all models were trained from scratch without using pre-trained weights from other datasets. The results of these experiments are shown in Table 6.

Table 6. Results of different models on the CrackSeg9k dataset.

Network Structure	MIoU (%)	Inference Time (s)	FLOPs (GFLOPs)
U-Net	80.69	0.013	124.01
PSPNet(Res50)	80.80	0.012	109.22
DeepLabV3(Res50)	80.22	0.015	165.57
BiseNetV2	78.85	0.002	7.513
DDRNet	75.39	0.003	11.122
HrSegNet-B48	79.26	0.002	5.036
CBU-FI Net	81.38	0.004	15.692

CBU-FI Net achieved an MIoU of 82.47% and an inference time of 0.021 s per image, setting a benchmark in crack detection for structures like masonry walls, ceramics, and pavement. It performed well on cracks ranging from 0.06 mm to 25 mm, demonstrating strong generalization across diverse surfaces. In contrast, U-Net and PSPNet struggled with finer cracks due to limitations in feature extraction. DeepLab-V3 (Res50) handled larger cracks better but was slower and more computationally demanding. Lightweight models like BiseNetV2, DDRNet, and HrSegNet-B48, while faster, sacrificed accuracy on narrower cracks. Their simpler architectures favored speed over precision, limiting their

effectiveness in detecting microcracks. CBU-FI Net, by balancing model complexity and detail retention, demonstrated superior accuracy and adaptability across a wide range of crack sizes and backgrounds.

3.2. Results of Ablation Experiments

3.2.1. Experimental Results on Composite Loss Parameters

We observed the impact of varying the weight of the adjustment coefficient on model performance under the same experimental conditions to investigate the setting of the adjustment coefficient in the proposed composite loss function. The experimental results are illustrated in Figure 6, where the horizontal axis represents the adjustment coefficient (α) for the Dice loss term in the composite loss function. The left vertical axis represents the Crack-IoU value, while the right vertical axis represents the pixel accuracy value.

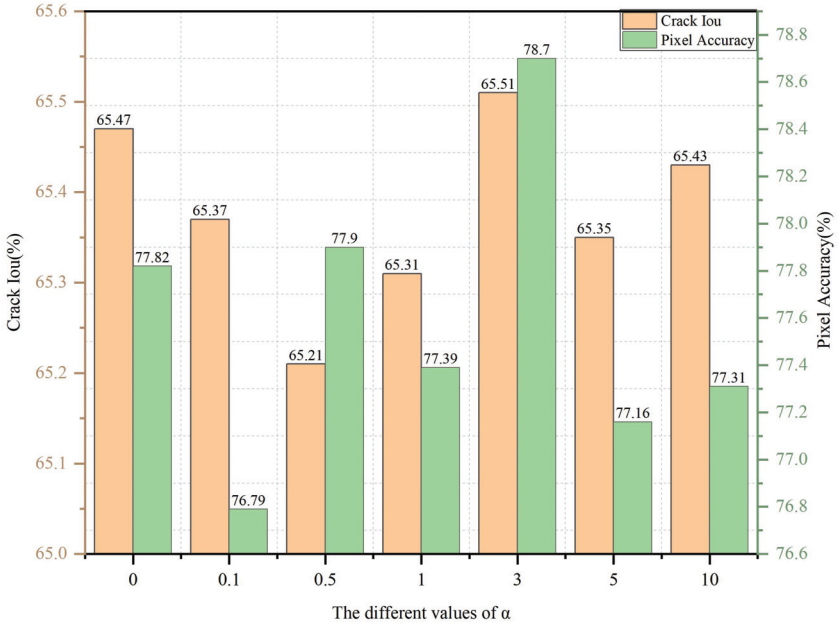


Figure 6. Results of different settings for the adjustment coefficient (α) in the composite loss function.

Based on our analysis, we can conclude the following. When the weight of the cross-entropy loss is relatively high, the model's accuracy decreases. This is because cross-entropy loss focuses more on the overall segmentation results and is less effective at capturing detailed features in microcrack detection. As the weight of the Dice loss increases, the model's performance improves. Under the given experimental conditions, setting the parameter weight to 3 (with a ratio of cross-entropy loss to Dice loss weight of 1:3) yields the optimal max Crack-IoU, indicating that the model achieves its best performance at this ratio. However, as the Dice loss weight increases further, the model's performance gradually declines slightly. This suggests that the model's performance is relatively insensitive to changes in the Dice loss, indicating a certain robustness in the selection of loss function weights.

3.2.2. Verification of Model Ablation Experiments

We conducted ablation experiments to further verify the effectiveness of the CBU-FI Net model in crack segmentation tasks. By incrementally adding various components to the model, we assessed their impact on performance. Table 7 summarizes the mean

intersection over union (MIoU), parameter count, and floating-point operations (FLOPs) for each ablation experiment configuration.

Table 7. Ablation results of each module of the CBU-FI Net model.

U-Net	ConvNeXt Block	Feature Injuction	Fuse	Cross-Entropy Loss	Dice Loss	MIoU (%)	Parameters (M)	FLOPs (GFLOPs)
√				√		80.28	29.06	811.64
√	√			√		81.68	5.055	27.34
√	√	√		√		82.10	5.430	52.54
√	√	√	√	√		82.45	7.696	102.88
√	√	√	√	√	√	82.47	7.696	102.88

Analysis of the ablation experiments clearly demonstrates the impact of each component on the performance enhancement of the CBU-FI Net model. Adding the ConvNeXt block resulted an increase in MIoU to 81.68%, while significantly reducing the parameter count to 5.055 M and FLOPs to 27.34 GFLOPs. This indicates that the ConvNeXt block effectively improves segmentation accuracy while reducing model complexity. The subsequent inclusion of the feature infusion module increased the MIoU to 82.10%, with a slight rise in parameter count to 5.430 M and FLOPs to 52.54 GFLOPs, highlighting the contribution of feature infusion to segmentation accuracy while maintaining a low computational complexity. The introduction of the Fuse module further raised the MIoU to 82.45%, with the parameter count and FLOPs increasing to 7.696 M and 102.88 GFLOPs, respectively, demonstrating the critical role of the Fuse module in feature fusion, significantly enhancing model performance. Finally, refining the composite loss function led to a slight MIoU increase to 82.47%, with no change in parameter count or FLOPs, indicating that the designed loss function further optimized model performance during the final fine-tuning stage, achieving the best MIoU. Overall, the results of the ablation experiments further validate the effectiveness of the introduced improvements.

3.3. Detection Visualization and Analysis

This section presents a qualitative analysis of the performance of CBU-FI Net and six other models across various datasets. Through visualization, we comprehensively evaluate each model's segmentation performance under different conditions. The qualitative analysis provides insights into how each model handles diverse types of cracks, varying background complexities, and changes in image quality, thereby enhancing our understanding of their robustness and generalization capabilities.

3.3.1. Visualization Analysis of the Poultry Egg Dataset

To evaluate the performance of each model in the microcrack detection task, we used samples characterized by lower clarity, the presence of texture interference, and finer cracks. The experimental results are illustrated in Figure 7, with red boxes indicating the areas of missed detection.

The original U-Net network exhibits relatively poor performance in crack segmentation, with some crack regions appearing fragmented and unable to be completely segmented. Additionally, the finer parts at the ends of cracks are not accurately segmented. This may be due to the structural limitations of the U-Net network in extracting microcrack features, leading to less detailed segmentation and a significant number of missed detections. Although other advanced models have shown improvements in microcrack detection, they still exhibit breakage, especially when crack features are not prominent, making it difficult to fully segment the crack regions. In contrast, the CBU-FI Net model proposed in this paper performs exceptionally well in crack segmentation tasks across various scales and shapes. Its overall detection results are closest to the ground truth, with significantly reduced breakage in crack segmentation, accurately segmenting the fine parts at the ends of cracks, thereby making the segmented crack details and boundaries clearer.

In summary, when comparing the results in the figures, it is evident that the improved algorithm proposed in this paper demonstrates outstanding segmentation performance in the poultry egg microcrack detection task. Compared to the original U-Net and other advanced models, the CBU-FI Net model effectively segments cracks. The segmentation visualization results indicate that this model has a significant advantage in handling microcracks of various scales and shapes, better capturing the detailed features of cracks and fully segmenting microcrack regions. This effectively reduces segmentation breakage and enhances the clarity of crack boundaries, making the shape of the cracks more realistic and accurate. This improvement has significant application value for microcrack detection tasks in the poultry egg industry.

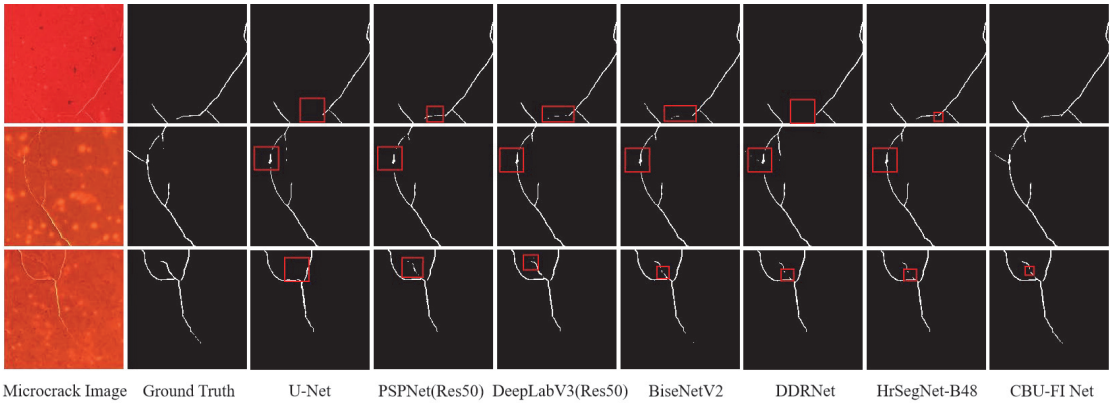


Figure 7. Visualization results of different models. The red boxes in the figure indicate missed detections.

3.3.2. Visualization Analysis of the CrackSeg9k Dataset

We selected five samples from the CrackSeg9k test set for visualization analysis to highlight the effectiveness of the proposed method. These samples encompass a range of crack scenarios, including asphalt pavement, concrete pavement, stone brick walls, and tile floors. By analyzing the crack segmentation results across these diverse scenarios, we can comprehensively assess the robustness and generalization capabilities of the proposed CBU-FI Net model in real-world applications.

The results indicate that in multi-scenario crack segmentation tasks, larger models perform slightly better than lightweight models, suggesting that computational limitations can affect model segmentation capabilities. When detecting pavement cracks, even though the materials differ, most models can segment the main shape of the cracks, since the crack features are distinctly different from the background texture. However, issues such as missed detections at the crack ends and crack breakage still occur. In contrast, when detecting cracks in tiles and stone bricks, the high similarity between the crack features and the background texture makes them harder to distinguish. Only PSPNet (Res50), DeepLab (Res50), and the proposed CBU-FI Net successfully detected the main cracks. Notably, PSPNet (Res50) and DeepLab (Res50) still exhibit crack breakage and missed ends. The proposed CBU-FI Net model significantly improves microcrack detection performance against low-contrast backgrounds through the design of feature infusion and feature fusion strategies. Compared to other models, CBU-FI Net excels in crack segmentation tasks for various scales and shapes, with detection results closest to the ground truth. This model notably reduces crack breakage and enhances the clarity of crack details and boundaries, demonstrating exceptional robustness and generalization capabilities in handling complex backgrounds and diverse crack features. In summary, the proposed CBU-FI Net model demonstrates superior segmentation performance across various crack scenarios, validating its effectiveness and practical value in crack segmentation tasks. Its design strategies not only improve detection accuracy but also significantly enhance the segmentation of crack

details and boundaries, providing a reliable solution for crack detection in real-world applications (Figure 8).

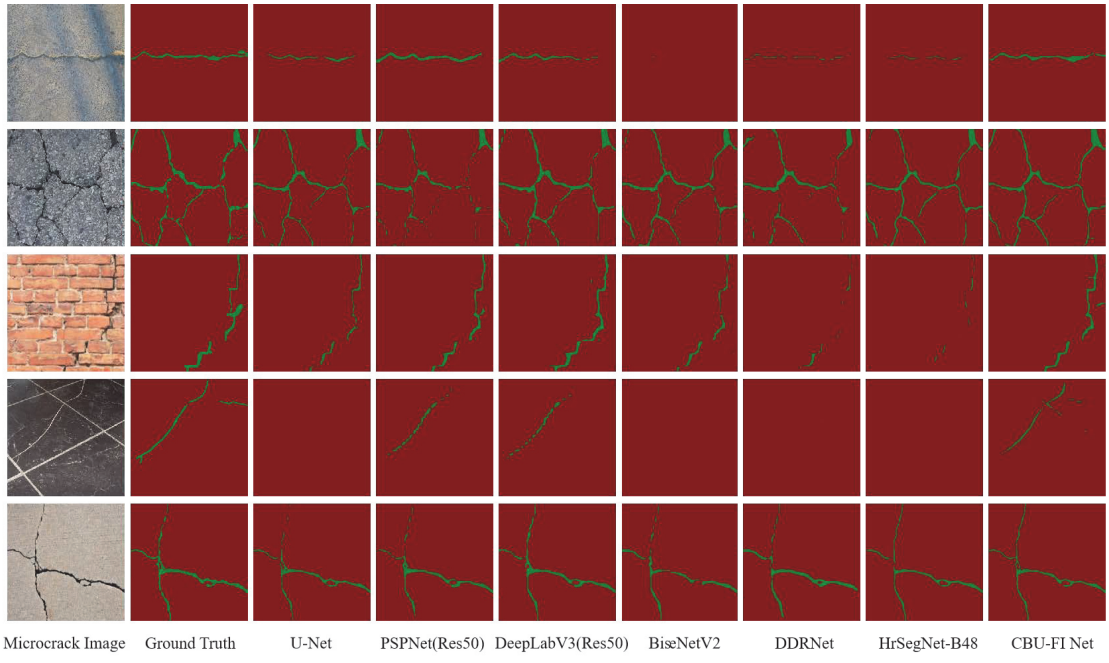


Figure 8. Visualization analysis of various models on the CrackSeg9k dataset.

4. Discussion

This study addresses the low efficiency and high cost of traditional poultry egg microcrack detection by proposing a real-time U-Net architecture based on feature infusion (CBU-FI Net). Using U-Net as the baseline model, we improved it to address the issues of large model parameters and slow image processing speed. The proposed model combines U-Net and ConvNeXt, enhancing the extraction of local detail information and global semantic information by injecting image edge features into the encoder and aggregating multi-scale features in the decoder. Additionally, a hybrid loss function based on cross-entropy loss and Dice loss was designed to further improve the segmentation performance of microcrack images, achieving real-time, high-precision, pixel-level microcrack detection. Compared to the original U-Net model, the improved model reduces the number of parameters and accelerates image processing speed while optimizing the model's focus on microcracks. This enhances the proportion of crack edge features within the overall semantic information, strengthening the model's ability to extract crack features.

The proposed real-time poultry egg microcrack detection model based on feature infusion achieves a real-time detection speed of 21 ms for input images with a resolution of 1024×1024. When detecting microcracks smaller than 20 μm, the model achieved an MIoU of 82.47%. On the benchmark CrackSeg9k dataset, it achieved an inference speed of 4 ms per image with a resolution of 400×400 and an MIoU of 81.38%, with a model parameter count of 7.696 M. Compared to state-of-the-art segmentation models, it achieves leading accuracy with fewer parameters. Using industrial microscopic measurement, the smallest detectable crack size is 3 μm. According to the visualization results presented in Section 3.3, although microcrack segmentation is relatively complete, there are still small parts of microcrack ends that are not fully detected. This may be due to the microcrack width being less than 3 μm or insufficient image resolution, resulting in the cracks occupying fewer

pixels in the image and making it difficult for the algorithm to accurately segment the complete outline of the cracks.

The high-efficiency microcrack detection system we propose can help to improve production efficiency on poultry egg production lines, reduce economic losses caused by defective eggs, and enhance agricultural production efficiency. It can positively impact the upgrading and development of the entire poultry egg industry.

5. Conclusions

This study proposes a real-time ConvNext-based U-Net with feature infusion for egg microcrack detection (CBU-FI Net). This model integrates the advantages of ConvNeXt's large convolution kernels and depthwise separable convolutions, addressing the issues of large parameter size and slow inference speed in the U-Net model. To tackle challenges in microcrack detection, such as narrow crack width and low contrast between crack ends and the background texture, the model infuses image edge features in the encoder and aggregates multi-scale features in the decoder, enhancing the extraction of both local detail information and global semantic information. Additionally, a composite loss function is constructed to address the imbalance between positive and negative samples. Using a single-point backlight source collection device, over 3400 graded poultry egg microcrack image patches were created for model training and validation.

Experimental results demonstrate that CBU-FI Net's parameter size is only one-third of that of the original U-Net, and the inference speed is 21 ms per image (1 million pixels). This model exhibits strong robustness and generalization ability, adapting to different types of cracks and complex background environments. On the public benchmark CrackSeg9k dataset, CBU-FI Net achieved an inference speed of 4 ms per 400×400 image and an MIoU of 81.38%. Additionally, on the poultry egg microcrack defect dataset, the model's Crack-IoU for microcracks (less than 20 μm) was 65.51%; for smaller cracks (less than 5 μm), the detection results show a Crack-IoU and MIoU of 60.76% and 80.22%, respectively, achieving real-time, high-precision real-microcrack detection. These results are superior to those achieved by the traditional U-Net model and other advanced semantic segmentation models, laying a foundation for online microcrack detection in poultry eggs based on semantic segmentation.

The proposed method for detecting microcracks in poultry eggs can significantly enhance production efficiency on egg processing lines, reducing economic losses associated with undetected defects. By improving the accuracy and reliability of crack detection, this method contributes to higher standards of food safety and quality control. Additionally, the method can be extended to other industries requiring precise defect detection, such as materials science, automotive manufacturing, and aerospace. As the technology continues to develop, it may find broader applications, driving innovation and efficiency across various sectors and supporting sustainable growth and advancement in these industries. In the future, we will explore the integration of discharge methods with imaging techniques to develop a multi-sensor fusion approach for the detection of microcracks in poultry eggs. By combining discharge signals with image data, we aim to enhance detection accuracy and robustness, thereby further improving the effectiveness of microcrack detection.

Author Contributions: Conceptualization, C.S. and C.Z. (Chun Zhang); methodology, C.S., C.Z. (Chun Zhang) and Y.L.; software, Y.L. and X.J.; validation, C.S., C.Z. (Chun Zhang), C.Z. (Changsheng Zhu), Y.L. and Y.M.; formal analysis, C.S., C.Z. (Chun Zhang), Y.L., X.J., Y.M., S.Y. and W.S.; investigation, C.S., Y.L., and C.Z. (Changsheng Zhu); resources, C.S. and Y.M.; data curation, Y.L., X.J. and W.S.; writing—original draft preparation, C.S., C.Z. (Chun Zhang) and Y.L.; writing—review and editing, C.S., Y.L., C.Z. (Changsheng Zhu), X.J. and W.S.; visualization, C.S., C.Z. (Chun Zhang), Y.L., C.Z. (Changsheng Zhu), Y.L., W.S. and Y.M.; supervision, C.S., C.Z. (Chun Zhang) and C.Z. (Changsheng Zhu). All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by the Shandong Province Science and Technology SME Innovation Capability Enhancement Project (No. 2023TSGC0576 and No. 2023TSGC0605) and the Tai'an Science and Technology Innovation Development Plan (No. 2021GX050 and No. 2020GX055).

Institutional Review Board Statement: Not applicable.

Data Availability Statement: The datasets presented in this article are not readily available because the data are part of an ongoing study. Requests to access the datasets should be directed to skd996523@sdust.edu.cn. The data derived from public-domain resources in this study are available in the CrackSeg9k dataset at https://doi.org/10.1007/978-3-031-25082-8_12 [27].

Acknowledgments: Special thanks to the Intelligent Perception and Cognition Laboratory at Shandong University of Science and Technology for providing the necessary facilities and resources to support this research.

Conflicts of Interest: The authors declare no conflicts of interest.

References

1. Brake, J.; Walsh, T.; Benton, C., Jr.; Petitte, J.; Meijerhof, R.; Penalva, G. Egg handling and storage. *Poult. Sci.* **1997**, *76*, 144–151. [CrossRef] [PubMed]
2. de Abreu Fernandes, E.; Litz, F.H. The eggshell and its commercial and production importance. In *Egg Innovations and Strategies for Improvements*; Elsevier: Amsterdam, The Netherlands, 2017; pp. 261–270.
3. Mazzuco, H.; Bertechini, A.G. Critical points on egg production: Causes, importance and incidence of eggshell breakage and defects. *Ciência e Agrotecnologia* **2014**, *38*, 7–14. [CrossRef]
4. Sun, L.; Zhang, P.; Feng, S.; Qiang, M.; Cai, J. Eggshell crack detection based on the transient impact analysis and cross-correlation method. *Curr. Res. Food Sci.* **2021**, *4*, 716–723. [CrossRef] [PubMed]
5. Sun, L.; Feng, S.; Chen, C.; Liu, X.; Cai, J. Identification of eggshell crack for hen egg and duck egg using correlation analysis based on acoustic resonance method. *J. Food Process. Eng.* **2020**, *43*, e13430. [CrossRef]
6. Bain, M.; MacLeod, N.; Thomson, R.; Hancock, J.W. Microcracks in eggs. *Poult. Sci.* **2006**, *85*, 2001–2008. [CrossRef]
7. Shi, C.; Wang, Y.; Zhang, C.; Yuan, J.; Cheng, Y.; Jia, B.; Zhu, C. Nondestructive Detection of Microcracks in Poultry Eggs Based on the Electrical Characteristics Model. *Agriculture* **2022**, *12*, 1137. [CrossRef]
8. Shi, C.; Cheng, Y.; Zhang, C.; Yuan, J.; Wang, Y.; Jiang, X.; Zhu, C. Wavelet Scattering Convolution Network-Based Detection Algorithm on Nondestructive Microcrack Electrical Signals of Eggs. *Agriculture* **2023**, *13*, 730. [CrossRef]
9. Purahong, B.; Chaowalittawin, V.; Krungseanmuang, W.; Sathaporn, P.; Anuwongpinit, T.; Lasakul, A. Crack Detection of Eggshell using Image Processing and Computer Vision. *J. Phys. Conf. Ser.* **2022**, *2261*, 012021. [CrossRef]
10. Guanjun, B.; Mimi, J.; Yi, X.; Shibo, C.; Qinghua, Y. Cracked egg recognition based on machine vision. *Comput. Electron. Agric.* **2019**, *158*, 159–166. [CrossRef]
11. Yao, K.; Sun, J.; Chen, C.; Xu, M.; Zhou, X.; Cao, Y.; Tian, Y. Non-destructive detection of egg qualities based on hyperspectral imaging. *J. Food Eng.* **2022**, *325*, 111024. [CrossRef]
12. Botta, B.; Gattam, S.S.R.; Datta, A.K. Eggshell crack detection using deep convolutional neural networks. *J. Food Eng.* **2022**, *315*, 110798. [CrossRef]
13. Tang, W.; Hu, J.; Wang, Q. High-throughput online visual detection method of cracked preserved eggs based on deep learning. *Appl. Sci.* **2022**, *12*, 952. [CrossRef]
14. Nasiri, A.; Omid, M.; Taheri-Garavand, A. An automatic sorting system for unwashed eggs using deep learning. *J. Food Eng.* **2020**, *283*, 110036. [CrossRef]
15. Xu, X.; Gao, Y.; Fu, C.; Qiu, J.; Zhang, W. Research on the Corn Stover Image Segmentation Method via an Unmanned Aerial Vehicle (UAV) and Improved U-Net Network. *Agriculture* **2024**, *14*, 217. [CrossRef]
16. Liu, C.; Su, J.; Wang, L.; Lu, S.; Li, L. LA-DeepLab V3+: A Novel Counting network for pigs. *Agriculture* **2022**, *12*, 284. [CrossRef]
17. Liu, Z.; Mao, H.; Wu, C.Y.; Feichtenhofer, C.; Darrell, T.; Xie, S. A convnet for the 2020s. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, New Orleans, LA, USA, 18–24 June 2022; pp. 11976–11986.
18. Han, Z.; Jian, M.; Wang, G.G. ConvUNeXt: An efficient convolution neural network for medical image segmentation. *Knowl.-Based Syst.* **2022**, *253*, 109512. [CrossRef]
19. He, Z.; Chen, W.; Zhang, J.; Wang, Y.H. Infrastructure crack segmentation: Boundary guidance method and benchmark dataset. *arXiv* **2023**, arXiv:2306.09196.
20. Qin, X.; Zhang, Z.; Huang, C.; Gao, C.; Dehghan, M.; Jagersand, M. Basnet: Boundary-aware salient object detection. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, Long Beach, CA, USA, 15–20 June 2019; pp. 7479–7489.
21. Alshawi, R.; Hoque, M.T.; Ferdaus, M.M.; Abdelguerfi, M.; Niles, K.; Prathak, K.; Tom, J.; Klein, J.; Mousa, M.; Lopez, J.J. Dual Attention U-Net with Feature Infusion: Pushing the Boundaries of Multiclass Defect Segmentation. *arXiv* **2023**, arXiv:2312.14053.
22. Zhao, H.; Shi, J.; Qi, X.; Wang, X.; Jia, J. Pyramid scene parsing network. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Honolulu, HI, USA, 21–26 July 2017; pp. 2881–2890.

23. Lin, G.; Milan, A.; Shen, C.; Reid, I. Refinenet: Multi-path refinement networks for high-resolution semantic segmentation. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Honolulu, HI, USA, 21–26 July 2017; pp. 1925–1934.
24. Milletari, F.; Navab, N.; Ahmadi, S.A. V-net: Fully convolutional neural networks for volumetric medical image segmentation. In Proceedings of the 2016 Fourth International Conference on 3D Vision (3DV), Stanford, CA, USA, 25–28 October 2016; pp. 565–571.
25. Yeung, M.; Sala, E.; Schönlieb, C.B.; Rundo, L. Unified focal loss: Generalising dice and cross entropy-based losses to handle class imbalanced medical image segmentation. *Comput. Med. Imaging Graph.* **2022**, *95*, 102026. [CrossRef]
26. Vinogradova, K.; Dibrov, A.; Myers, G. Towards Interpretable Semantic Segmentation via Gradient-Weighted Class Activation Mapping (Student Abstract). *Proc. AAAI Conf. Artif. Intell.* **2020**, *34*, 13943–13944. [CrossRef]
27. Kulkarni, S.; Singh, S.; Balakrishnan, D.; Sharma, S.; Devunuri, S.; Korlapati, S.C.R. CrackSeg9k: A collection and benchmark for crack segmentation datasets and frameworks. In Proceedings of the European Conference on Computer Vision, Tel Aviv, Israel, 23–27 October 2022; Springer: Berlin/Heidelberg, Germany, 2022; pp. 179–195.
28. Ronneberger, O.; Fischer, P.; Brox, T. U-net: Convolutional networks for biomedical image segmentation. In Proceedings of the Medical Image Computing and Computer-Assisted Intervention–MICCAI 2015: 18th International Conference, Munich, Germany, 5–9 October 2015; proceedings, part III 18; Springer: Berlin/Heidelberg, Germany, 2015; pp. 234–241.
29. Chen, L.C.; Zhu, Y.; Papandreou, G.; Schroff, F.; Adam, H. Encoder-decoder with atrous separable convolution for semantic image segmentation. In Proceedings of the European Conference on Computer Vision (ECCV), Munich, Germany, 8–14 September 2018; pp. 801–818.
30. Yu, C.; Gao, C.; Wang, J.; Yu, G.; Shen, C.; Sang, N. Bisenet v2: Bilateral network with guided aggregation for real-time semantic segmentation. *Int. J. Comput. Vis.* **2021**, *129*, 3051–3068. [CrossRef]
31. Hong, Y.; Pan, H.; Sun, W.; Jia, Y. Deep dual-resolution networks for real-time and accurate semantic segmentation of road scenes. *arXiv* **2021**, arXiv:2101.06085.
32. Li, Y.; Ma, R.; Liu, H.; Cheng, G. HrSegNet: Real-time High-Resolution Neural Network with Semantic Guidance for Crack Segmentation. *arXiv* **2023**, arXiv:2307.00270.
33. Shi, C.; Li, Y.; Cheng, Y.; Wang, Y.; Zhu, C.; Wang, K.; Zhang, C. Detection of Microcrack in Eggs Based on Improved U-Net. In Proceedings of the 2023 IEEE 9th International Conference on Cloud Computing and Intelligent Systems (CCIS), Dali, China, 12–13 August 2023; pp. 409–413.

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.

agriculture

MDPI

Article

A High-Accuracy Contour Segmentation and Reconstruction of a Dense Cluster of Mushrooms Based on Improved SOLOv2

Shuzhen Yang [1,*], Jingmin Zhang [1] and Jin Yuan [2]

[1] School of Intelligent Manufacturing and Control Engineering, Shanghai Polytechnic University, Shanghai 201209, China
[2] College of Mechanical and Electronic Engineering, Shandong Agricultural University, Taian 271018, China; jyuan@sdau.edu.cn
* Correspondence: szyang@sspu.edu.cn; Tel.: +86-21-5021-6899

Abstract: This study addresses challenges related to imprecise edge segmentation and low center point accuracy, particularly when mushrooms are heavily occluded or deformed within dense clusters. A high-precision mushroom contour segmentation algorithm is proposed that builds upon the improved SOLOv2, along with a contour reconstruction method using instance segmentation masks. The enhanced segmentation algorithm, PR-SOLOv2, incorporates the PointRend module during the up-sampling stage, introducing fine features and enhancing segmentation details. This addresses the difficulty of accurately segmenting densely overlapping mushrooms. Furthermore, a contour reconstruction method based on the PR-SOLOv2 instance segmentation mask is presented. This approach accurately segments mushrooms, extracts individual mushroom masks and their contour data, and classifies reconstruction contours based on average curvature and length. Regular contours are fitted using least-squares ellipses, while irregular ones are reconstructed by extracting the longest sub-contour from the original irregular contour based on its corners. Experimental results demonstrate strong generalization and superior performance in contour segmentation and reconstruction, particularly for densely clustered mushrooms in complex environments. The proposed approach achieves a 93.04% segmentation accuracy and a 98.13% successful segmentation rate, surpassing Mask RCNN and YOLACT by approximately 10%. The center point positioning accuracy of mushrooms is 0.3%. This method better meets the high positioning requirements for efficient and non-destructive picking of densely clustered mushrooms.

Keywords: contour reconstruction; densely overlapping mushrooms; high-precision contour segmentation; image segmentation; SOLOV2

Citation: Yang, S.; Zhang, J.; Yuan, J. A High-Accuracy Contour Segmentation and Reconstruction of a Dense Cluster of Mushrooms Based on Improved SOLOv2. *Agriculture* **2024**, *14*, 1646. https://doi.org/10.3390/agriculture14091646

Academic Editor: Antonio Paz-González

Received: 10 August 2024
Revised: 13 September 2024
Accepted: 18 September 2024
Published: 20 September 2024



Copyright: © 2024 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (https://creativecommons.org/licenses/by/4.0/).

1. Introduction

Mushrooms, recognized as one of the edible fungi with the most extensive cultivated areas and found in some of the world's largest cultivation countries, exhibit a significant output. In China alone, they boast an annual production of nearly 2 million tonnes. Currently, the mainstream production mode of mushrooms in developed countries is factory production, and most of the processes have realized mechanization and automation. Only picking still relies on manual labor [1,2]. Due to the decline in the rural labor force, manual mushroom picking is experiencing a significant shortage of labor, posing a serious obstacle to the industry's continued growth. Scholars have researched the autonomous picking technology of mushrooms and achieved phased development. The relevant research has developed from the exploration stage [3,4], limited to the laboratory under ideal conditions, to the field experiment test stage [5,6] in the actual factory mushroom house. However, due to the heterogeneity of the realistic growing environment of mushroom houses and the randomness of fruiting body growth, it is challenging to achieve efficient and nondestructive picking, especially for densely overlapping mushrooms [7]. Clusters

of mushrooms has become the main reason that prevents the practical application of mushroom havesting robots. Therefore, identifying and locating densely overlapping mushrooms with high precision is one of the keys to realizing more efficient, nondestructive mushroom picking.

Mushrooms are a kind of natural high-cluster crop. It is effortless for fruiting bodies to grow together densely and overlappingly. Due to the limited space for dense growth, the mushroom shape can easily vary from round to oval or irregular, or vertical to inclined. Many mushrooms are seriously overlapping, which makes it difficult to segment and reconstruct the mushroom contour with high accuracy. For example, since densely growing mushrooms are easily squeezed and stacked, the contour segmentation between them will become complex, and it is not easy to obtain accurate edge contours of mushrooms, as shown in Figure 1, Mushrooms 1, 2, 3, and 4. For another example, densely growing mushrooms are prone to appearing heavily obscured, such as mushrooms 9 and 10 in the figure. Their shapes are no longer round or elliptical but are missing and irregular. It is often difficult to restore the contour based on these missing or irregular shapes to the actual contour, such as contour 9. The deviation between the generated and accurate contours is significant, resulting in a considerable variation between the identified and actual mushroom center points. In addition, a few small mushrooms are easily overlooked, such as Mushroom 5, 6, 7, and 8. Therefore, to achieve efficient and lossless picking of densely overlapping mushrooms, it is necessary to study further and improve the segmentation and contour reconstruction algorithms of densely overlapping mushrooms to obtain high-precision positioning.

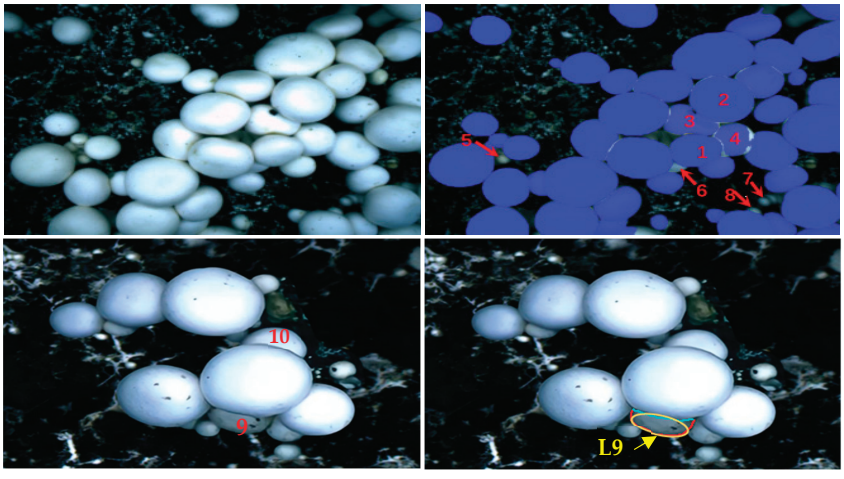


Figure 1. Main problems in segmentation and localization of densely overlapping mushrooms.

The existing system faces several challenges in accurately segmenting and reconstructing mushroom contours:

Dense Overlapping Growth: Generally, mushrooms grow in clusters with densely overlapping caps. Conventional segmentation methods encounter difficulties in distinguishing individual mushrooms within these clusters, resulting in inaccurate boundary delineation.

Deformation and Irregular Shapes: Due to mutual adhesion, extrusion, or varying growth conditions, mushrooms can exhibit irregular shapes and deformities. Most of the methods, fitting with circles only, are not very accurate in fitting the contours of irregular shapes and deformities.

Incomplete and Occluded Contours: When mushrooms are partially covered by other mushrooms, the segmentation process may yield incomplete or occluded contours, making it challenging to obtain an accurate representation of each mushroom's shape.

Variable Lighting and Impurities: Non-uniform lighting conditions and the presence of surface impurities, such as soil and lesions, can adversely affect the quality of mushroom images, leading to additional segmentation difficulties.

To address these problems, the proposed work makes significant contributions:

(1) The PR-SOLOv2 instance segmentation model is introduced that leverages PointRend to improve segmentation accuracy. This model effectively handles dense overlapping mushrooms and enhances boundary accuracy, achieving a segmentation accuracy of 93.04%.
(2) The work proposes a novel contour reconstruction method based on the curvature and length of mask contours obtained through instance segmentation. This approach enables the reconstruction of mushroom contours with high precision.
(3) The work introduces a least-squares ellipse fitting technique for mushrooms with regular shapes. For irregularly shaped or occluded mushrooms, a corner-based segmentation method is applied to accurately reconstruct contours. This classification and fitting approach significantly improves the quality of contour reconstruction.
(4) The proposed work demonstrates a successful segmentation rate of 98.92%. This high success rate shows the effectiveness of the proposed methods.

The manuscript is structured as follows: Section 2 reviews existing research and methodologies related to mushroom segmentation and contour reconstruction. Section 3 elucidates mushroom instance segmentation dataset construction. Section 4 introduces the PR-SOLOv2 instance segmentation model. The contour construction of mushrooms using the proposed approach is presented in Section 5. It explains the classification and fitting approach based on curvature and length. Section 6 presents the experimental results of the proposed methods and compares them to other segmentation algorithms, highlighting the accuracy and effectiveness of PR-SOLOv2. Section 7 summarizes the major contributions of the research, emphasizing the improvements in mushroom segmentation and contour reconstruction.

2. Related Work

In 1968, American researchers Schertz et al. initially suggested the utilization of image processing technology for citrus detection and picking [8]. Subsequently, computer vision started to be extensively employed in detecting fruits and vegetables. The research on identifying and positioning single fruits and vegetables also started with traditional visual technology and developed rapidly [9–12]. Several researchers [10,13] adopted the machine learning method to detect and locate apples and citrus fruits by extracting the features of the target in the image. However, it was not easy to simultaneously ensure high precision, high efficiency, and detection robustness.

As deep learning technology has advanced, the application of convolutional neural network (CNN)-based architectures and their variations has progressively been employed in the detection of fruits and vegetables, demonstrating notable advantages [14]. Researchers have suggested employing enhanced target detection algorithms, such as the Faster RCNN (Faster Region-based Convolutional Neural Network) algorithm, the Single Shot MultiBox Detector (SSD) algorithm, and You Only Look Once (YOLO), to identify the locations of strawberries, apples, and other fruits. These algorithms have demonstrated a recognition success rate exceeding 83%. It has been proven that it is quite feasible to use the target detection class deep learning algorithm to calculate the quantity and judge the quality of plant phenotypes [15,16]. In recent years, due to the improvement of people's requirements for target positioning and the need to distinguish different individuals in the map, instance segmentation methods have emerged as the times require. Among these, the Mask Region-based Convolutional Neural Network (Mask RCNN), a convolutional neural network based on regions, has also been explored for its application in recognizing the picking of grapes, oranges, strawberries, apples, and other fruits and vegetables, and, in different light intensities, multi-fruit adhesion and overlapping. The success rate of fruit recognition under complex growth conditions such as tree occlusion can reach more

than 90%, which shows that the case segmentation method has significant robustness in agricultural fruit detection [17–20].

In the exploration of recognizing and locating edible fungi, researchers have experimented with traditional image processing methods. Morphological target detection algorithms, including threshold methods based on grayscale [21], sequential scanning algorithms [22], watershed algorithms based on angular density features [23], segmentation and recognition methods based on edge grayscale gradient features, and contour grouping [24] have been employed to identify and pinpoint mushroom contours. As a result, the recognition success rate has risen from approximately 80% to around 95%. However, the generalization ability of traditional visual methods is poor, and they are vulnerable to illumination and other objective factors. This kind of method's segmentation and positioning effect for mushrooms with sparse or general growth density is acceptable. However, the segmentation and positioning effect for mushrooms with complex adhesion and high-density overlap will be significantly reduced, which cannot meet the high-precision positioning requirements for efficient and nondestructive picking of densely overlapping mushrooms.

The advancement of machine learning has extensively promoted the research pro-gress in recognizing edible fungi [25]. However, it is used in the classification of edible fungi varieties or growth stages [26–28]. Deep learning models are becoming a new trend in edible fungus detection, which further optimize the effectiveness of edible fungus identification, classification, and localization. At present, the significance of deep learning technology in the field of edible fungi is more focused on classification research [29–33], and only a few scholars have carried out in-depth studies on the segmentation and localization of mushroom fruiting bodies by using this technology. Lee et al. [34] established a faster RCNN target detection model to identify mushrooms and obtained three-dimensional point cloud data through a depth camera to segment a single mushroom. The identification accuracy of this algorithm was 70.93%. The oyster mushroom harvesting robot designed by Rong et al. [35] adopted the SSD target detection algorithm improved by MobileNet-v2 to identify oyster mushrooms, with a 95% success rate and improved detection speed.

Zheng et al. [36] used two detection network models, YOLOv3 and YOLOv4, to de-tect the head of Flammulina and found that the test time of the YOLOv4 algorithm was 0.8 s and the test accuracy was 81.54%, which was superior to YOLOv3. Lu et al. [37] applied the YOLOv3 algorithm for the identification of *Agaricaria bisporus* and introduced a positioning correction method to enhance the results. They devised the fractional penalty algorithm to compute the circle diameter of the mushroom cap, assess the fruiting body's growth rate, and estimate harvest time, thereby enhancing the efficiency of mushroom growth management. Nevertheless, the accuracy of the algorithm can be influenced by the clarity of the contour in the mushroom region. The contour reconstruction method based on the circle has a specific deviation for the edge fitting of rare mushrooms with severe adhesion and extrusion deformation.

Missourian [38] focused on image classification and segmentation in the context of a harvesting robot's 3D vision system. This work employs a Support Vector Machine (SVM) to identify mushrooms as either class one or class two. Image segmentation is addressed through a multi-step process using HSV color space, image gradient information, and the Hough transform to locate individual mushrooms in the XY plane. Sert and Okumus [39] studied segmenting mushrooms in images and measuring their cap width. They make use of a customized K-Means clustering method, enhancing classical K-Means for better segmentation. Image preprocessing includes filtering and histogram balancing, followed by segmentation. Soomro et al. [40] presented a two-stage image segmentation method for intensity inhomogeneous images. The first stage combines global intensity and geodesic edge terms, producing a rough segmentation. The second stage refines the segmentation using local intensity and geodesic edge terms. They use image gradient information and a Gaussian kernel for regularization in their energy function.

Pchitskaya and Bezprozvanny [41] examined the shapes of dendritic spines, emphasizing their dynamic nature and advocating for clusterization approaches over classifi-

cation. They explore algorithms that rely on clusterization to enable new levels of spine shape analysis. Ref. [42] presented a method for detecting and measuring the diameter of mushroom caps using depth image processing. They proposed a novel approach based on the three-dimensional structure of mushrooms to segment them from the background and measure their diameters using the Hough transform. Arjun et al. [43] analyzed visual cues for white button mushrooms using image processing tools. It assesses chromatic and morphological characteristics and predicts the onset of mushroom discoloration through hyperspectral image analysis.

Baisa and Al-Diri [44] presented an algorithm for detecting, localizing, and estimat-ing the 3D pose of mushrooms using RGB-D data from a consumer RGB-D sensor. It combines RGB and depth information for these purposes and employs advanced techniques for 3D pose estimation. Zhao et al. [45] proposed a white measurement method for *Agaricus bisporus* based on image analysis. It involves the construction of an imaging system, color calibration, and the application of CIE Ganz whiteness formula to assess the whiteness of the mushrooms. Retsinas et al. [46] concentrated on developing a vision module for 3D pose estimation of mushrooms in an industrial mushroom farm. It uses multiple RealSense active-stereo cameras and employs a novel pipeline for mushroom instance segmentation and template matching.

Zhang et al. [47] investigated more than 62 articles related to microorganism biovolume measurement using digital image analysis methods. It tracks the development of this approach since the 1980s. Chen et al. [48] introduced an improved YOLOv5s algorithm for accurate *Agaricus bisporus* detection. The algorithm combines a deep learning network with the Convolutional Block Attention Module (CBAM) to enhance detection accuracy. Chen et al. [49] focused on accurately recognizing *Agaricus bisporus* in complex backgrounds by proposing a watershed-based segmentation recognition algorithm. It includes various image processing techniques for accurate mushroom segmentation and recognition. Table 1 illustrates a detailed summary of the existing works related to the proposed work.

Table 1. Summary of the state-of-the-art works.

Author	Approach	Major Findings	Segmentation Success Rate	Mean Time
Masoudian [38]	SVM for image classification, multi-step segmentation	This approach is not well suited for large datasets.	88.57%	0.49 s
Soomro et al. [40]	Two-stage segmentation for intensity inhomogeneous images	The two-stage segmentation technique overcomes issues related to the initial contour, making it suitable for inhomogeneous image segmentation.	97.99%	1.41 s
Ji et al. [42]	Depth image processing for mushroom diameter measurement	The method detects mushrooms and provides essential information for robotic selective harvesting, improving production quality and efficiency.	92.37%	0.5 s
Baisa and Al-Diri [44]	RGB-D data for mushroom detection and 3D pose estimation	The algorithm provides a robust method for mushroom detection, localization, and 3D pose estimation, which can be valuable for robotic picking applications.	98.99%	-
Retsinas et al. [46]	Vision module for 3D pose estimation using Intel RealSense cameras	The method aims to accurately detect the 3D pose of mushrooms without relying on 3D annotation data, which can facilitate mushroom harvesting by robotic systems.	96.31%	3 s
Zhang et al. [47]	Review of microorganism biovolume measurement methods	The work has research significance and application value, offering insights into microorganism biovolume measurement using digital image analysis methods.	-	-

Table 1. *Cont.*

Author	Approach	Major Findings	Segmentation Success Rate	Mean Time
Chen et al. [48]	Improved YOLOv5s algorithm for *Agaricus bisporus* detection	The improved algorithm shows convincing recognition accuracy, robustness, and low error rates in *Agaricus bisporus* detection.	98%	0.018 s
Chen et al. [49]	Watershed-based segmentation for *Agaricus bisporus* recognition	The proposed algorithm demonstrates recognition of *Agaricus bisporus* with low error rates and real-time processing capabilities.	95.7%	0.706
Tillett et al. [21]	Algorithm for locating mushrooms in growing beds	The algorithm is designed for use in robotic harvesting systems and provides a basis for identifying and tracking mushrooms in a growing bed.	93%	-
Yang et al. [23]	Corner density-based localization algorithm for overlapping mushrooms	The algorithm demonstrates a success rate of 86.3% in locating overlapping mushrooms, providing an efficient solution for mushroom localization.	96.37%	0.712 s
Yang et al. [24]	Improved segmentation recognition algorithm for overlapping *Agaricus bisporus*	The algorithm achieves over 90% recognition rate for *Agaricus bisporus* in overlapping situations, making it suitable for complex planting environments.	97.25%	0.212 s
Lu and Liaw [37]	CNN-based mushroom cap diameter measurement	The algorithm provides an estimation of mushroom cap diameters and outperforms the Circle Hough Transform.	82.7%	-

To sum up, traditional image recognition methods are still widely used in mushroom recognition and positioning, and the recognition effect has been continuously improved. However, due to the influence of model generalization, there is little room for further improvement. In contrast, the algorithm based on depth learning dramatically improves the accuracy and generalization ability of the recognition algorithm. Currently, the mainstream methods are YOLO, faster RCNN, and other target detection methods. Recently, Mask RCNN has begun to emerge in recognition of fruits and vegetables and has achieved better results. Yin et al. [50,51] also attempted to use instance segmentation to segment caps of *Oudemansiella raphanipes* and *Agrocybe cylindracea*. Moreover, instance segmentation proves especially effective in detecting fruits in complex growth states, including varying light intensities and instances of multi-fruit adhesion, overlap, and occlusion. The instance segmentation algorithm will become the next research hotspot in the recognition and positioning of fruits, vegetables, and edible fungi during picking.

However, two-stage algorithms, which normally have low efficiency, are mainly used at present. In contrast, the single-stage instance segmentation algorithm SOLO (Segmenting Objects by Locations) was simplified compared with the two-stage method of first detection and then segmentation; it claims a more straightforward design to achieve greater efficiency and better segmentation performance [52,53]. Furthermore, the improved version of the SOLO model, SOLOv2, uses dynamic convolution to enhance the quality of the mask segmentation and the model's efficiency [54]. In SOLOv2, the category branch and mask branch are integrated, and the matrix non-maximum suppression is proposed to inhibit the repetitive prediction after multi-scale feature extraction of the input image by Full Convolutional Networks (FCN), thus realizing faster mask processing [55]. Therefore, the SOLOv2 network structure is used as the segmentation framework for more and more tasks requiring efficient and accurate segmentation between the object and the background. For instance, Li et al. [56] utilized the SOLOv2 algorithm to segment the perimeter of the offshore area in the oilfield operation site and then took the segmentation result as the key to detecting the intrusion of the offshore perimeter area. The accuracy rate of the designed model was up to 94.7%; Ji et al. [57] used the method based on the SOLOv2 network and

point cloud cavity characteristics to automatically score the rumen filling of dairy cows. The average segmentation accuracy on the test set images was 86.29%, showing robustness to individual differences.

Therefore, given the shortcomings of low segmentation and positioning accuracy in visual recognition of densely overlapping mushrooms, this paper proposes a novel high-accuracy positioning method for densely overlapping mushrooms based on the instance segmentation model SOLOv2. This method has the following characteristics: (1) It solves the fuzzy problem of densely overlapping mushroom edge segmentation; (2) the most extended contour classification reconstruction algorithm based on the mask contour of instance segmentation can better eliminate the interference sub-contour in the irregular contour that is seriously occluded and extruded, and effectively improves the contour reconstruction effect and mushroom center point positioning accuracy through classification fitting.

3. Mushroom Instance Segmentation Dataset Construction

This paper takes mushrooms as the research object and collects images of mushrooms growing in the mushroom house of Shanghai Lianzhong Edible Fungus Professional Cooperative to construct the dataset. Mushrooms are raised in factory production environments and standard culture racks. The culture racks are 31 m long, 1440 mm wide, and 6 storeys high, with the storeys spaced 600 mm apart. The temperature, humidity, carbon dioxide concentration, and light in the mushroom room are intelligently controlled by an intelligent control system under conditions suitable for the growth of mushrooms, as shown in Figure 2. Three MV-GE300C-T GigE color industrial cameras, with a focal length of 4 mm, a pixel depth of 10 bits of colour and image resolution of 1280*720 is used, were installed transversely on the picking robot to traverse each layer of mushrooms and capture mushroom images.



Figure 2. Field image of mushroom image acquisition.

Mushrooms quickly form dense clusters, especially the mushrooms in the early picking stage, whose growth is particularly lush, resulting in mutual extrusion, shielding, staggering, and other phenomena among mushrooms. In addition, the fruiting bodies of mushrooms may appear on the surface of scales during the growth process or the browning phenomenon due to poor anti-browning ability, thus affecting the shape of mushrooms. Therefore, the construction of the data set should fully consider these differences in mushroom growth shape and posture. In this paper, a total of 480 images of eight typical forms of mushrooms were collected, and 60 shots were gathered for each category, as shown in Figure 3.

This paper also considers the images under different light source conditions when constructing the dataset. Specific composition (Figure 4): 160 images with uniform light and high brightness collected under the top light source, 160 pictures with uneven light under the side light source, and 160 images with low brightness under the dark light source.

The images in this paper are labeled with labelme. To ensure that the labeled polygons try to fit the real contours of the mushrooms, the number of basic labeling points for each mushroom example is stipulated, following the definition of a COCO dataset for s, m, l to divide the mushroom model: s-type mushrooms at least 10 labeling points, m-type at

least 20 labeling points, and l-type mushrooms at least 25 points, to improve the degree of labeling and fit to the real contours of the mushrooms. After labeling all the initial images to obtain the initial data set, we used data enhancement to process the images to improve the generalization ability and robustness of the model. Specific image data enhancement methods included: image brightness change, horizontal flipping of the image, vertical mirroring, and random rotation. Finally, 1920 mushroom images obtained after performing the image enhancement process were randomly divided into training and validation sets in a ratio of 7:3.

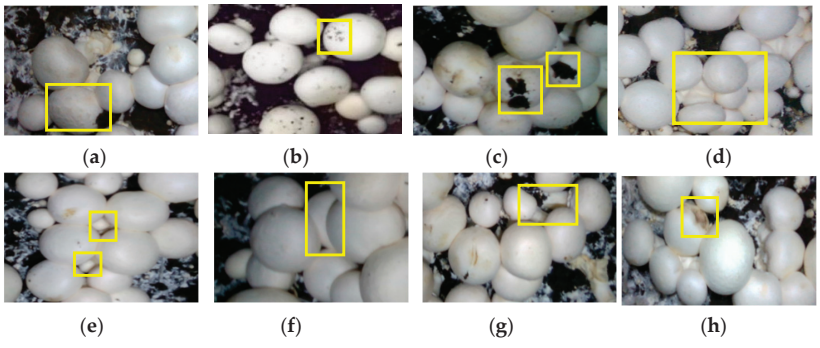


Figure 3. Datasets of different forms of mushrooms. (**a**) With scales on caps; (**b**) With brown spots on caps; (**c**) With soil on caps; (**d**) Serious interleaving, clumping, and adhesion; (**e**) Large differences in size and height; (**f**) Adhesion and extrusion deformation between mushrooms; (**g**) Heavily tilted to expose its stalk; (**h**) With mechanical damage on caps.

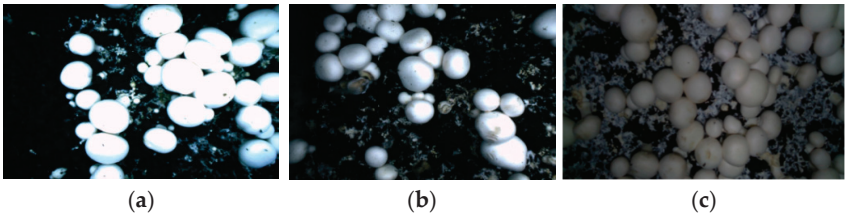


Figure 4. Datasets for different light sources: (**a**) Top light source; (**b**) Side light source; (**c**) Dark light source.

4. Research on High-Accuracy Mushroom Segmentation

The SOLOv2 algorithm, shown in Figure 5, was used to segment the mushrooms first.

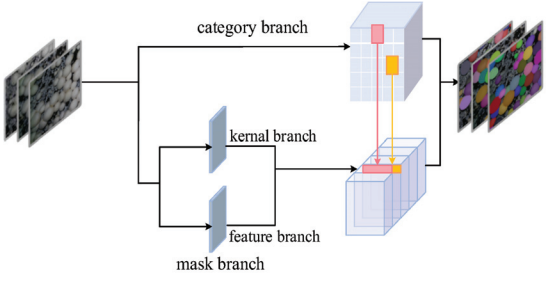
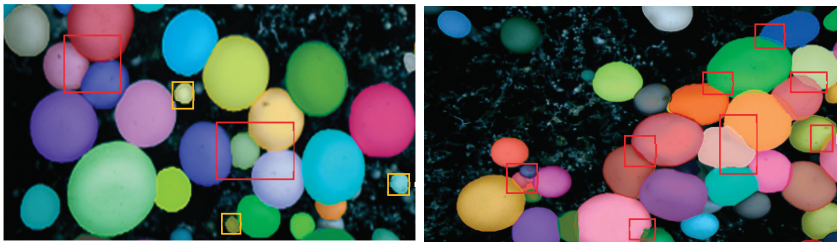


Figure 5. SOLOv2 network structure.

The segmentation results are as follows: for the mushrooms with a general aggregation degree and good growth form, the segmentation effect is good. The recognition rate of

the mushrooms in Figure 6a is high, especially since some tiny mushrooms are segmented. The mask boundary fits well with the contour of the actual mushroom. However, the segmentation effect will be affected for mushrooms with a high aggregation degree and complex growth morphology, resulting in a rough mask and inaccurate segmentation of mushroom edges, as shown in Figure 6b. Especially in the densely overlapping area of mushrooms, the positioning accuracy of the mushroom picking center will be affected. Therefore, SOLOv2 needs to be improved to obtain more precise and accurate segmentation of mask edges.



(**a**) (**b**)

Figure 6. Result of SOLOv2 segmented mushrooms: (**a**) Ordinary aggregation; (**b**) Complex polymerization.

The PointRend neural network can optimize object edge segmentation, treat segmentation of an image as a rendering problem, and perform the prediction of image segmentation based on an adaptive point selection iterative subdivision algorithm [58]. The structure of the network is shown in Figure 7. A lightweight prediction head generates a rough prediction mask, and bilinear interpolation is utilized for upsampling the rough prediction mask to obtain a prediction mask with twice the resolution. N points distributed on the edge are selected, and point-by-point features are constructed on the points chosen by combining fine-grained features and coarse prediction features. Finally, a simple multi-layer perceptron is used to predict these points. All points share the weight of the multi-layer perceptron. The multi-layer perceptron can be trained by standard task-specific segmental losses and finally output the classification label of each point. This is iterated until the mask at the desired resolution is upsampled. Therefore, this paper uses the PointRend network structure, which can solve the edge segmentation accuracy that is not high enough to improve the SOLOv2 network. In SOLOv2, to obtain the semantic information of the segmentation target, multiple convolutions and pooling operations on the original image will reduce the image resolution. Then, if the image with high semantic and low key is directly upsampled to be consistent with the size of the original image, information in specific areas will be lost, resulting in rough segmentation results. In addition, the method of predicting each pixel in a regular grid and inferring the instance category to which it belongs has the conditions of over-sampling and under-sampling. In particular, too many sampling points in low-frequency areas belonging to the same object in the image leads to over-sampling, and the under-sampling of such high-frequency regions close to the object boundary in the picture makes the segmented border too smooth and unreal. This will inevitably lead to rough segmentation results, especially at the edge of the segmentation object.

Therefore, this paper combines the SOLOv2 network structure with the PointRend neural network module to build an improved SOLOv2 algorithm, which is called the PR-SOLOv2 algorithm in this paper. Its network structure is shown in Figure 8. After the Fully Convolution Network (FCN) of SOLOv2's feature extraction backbone network, the PointRend module is introduced before the mask prediction branch to improve the image up-sampling method. To solve the information loss in local areas and render fine segmentation details, the delicate features of the second layer of the FCN network are

introduced, combined with the rough features of the fourth layer, and sent to the taught PointRend module.

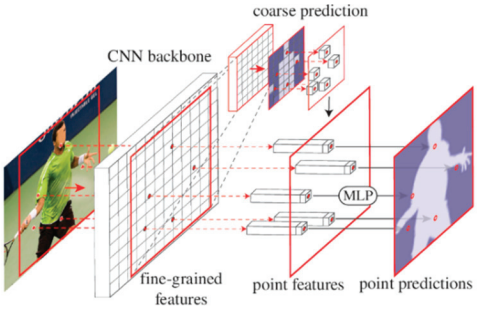


Figure 7. PointRend network structure.

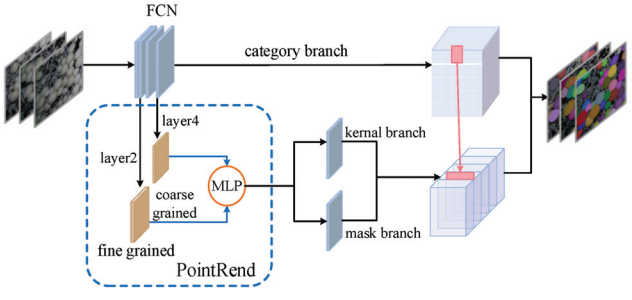


Figure 8. PR-SOLOv2 network structure.

The top-N feature pixels are extracted adaptively from the feature map of low spatial resolution and the corresponding fine feature map of high spatial resolution. The multi-layer perceptron is used to iterate and optimize each point continuously, and the instance is connected to the segmentation prediction branch to predict the instance mask. By this method, low-level and high-level features are combined layer by layer. The output image is iteratively rendered from coarse–fine to coarse–fine to solve the boundary inauthenticity caused by the over-sampling of the smooth region and under-sampling of the boundary region, which can improve the misjudgment of boundary pixels caused by the original network structure and improve the accuracy of the image segmentation edge.

The model's total loss function L is composed of the loss L_{cate} generated by instance classification and the loss L_{Mask} generated by mask prediction. λ is the coefficient that balances the two losses, and the calculation is shown in Equation (1).

$$L = L_{\text{cate}} + \lambda L_{\text{Mask}} \tag{1}$$

$$L_{\text{Mask}} = \frac{1}{N_{\text{pos}}} \sum 1\left\{P^*_{i,j} > 0\right\}^{d_{\text{Mask}}}, i, j \in M \tag{2}$$

$$d_{\text{Mask}} = 1 - D(p,q) = 1 - \frac{2\sum_{x,y}\left(p_{x,y} \bullet q_{x,y}\right)}{\sum_{x,y} p^2_{x,y} + \sum_{x,y} q^2_{x,y}} \tag{3}$$

where L_{cate} is the Focal Loss function for classification [59] for the calculation of L_{Mask} (Equation (2)). S is the number of grid units divided into the input image. i and j respectively represent the subscripts of indexes from left to right and from top to bottom on $M \times M$ grids on the input image, traversing all grids on the picture; Npos is the number of positive samples; P^* represents the instance category. $1\left\{P^*_{i,j} > 0\right\}$ is the indicator function; if $P^*_{i,j} > 0$,

the indicator function value is 1, otherwise it is 0; d_{Mask} is the Dice Loss function, and its definition is shown in Equation (3), where D is the Dice coefficient, which is used to measure the similarity of two sets. p and q represent the predicted Dice Loss mask and the natural mask, respectively, and $p_{x,y}$ and $q_{x,y}$ are the pixel values of the point (x, y) of the predicted mask and the natural mask on the image, respectively.

5. Contour Reconstruction of Mushroom

In the process of mushroom growth, due to its natural cluster growth characteristics, it is common for mushrooms to overlap and extrude each other, leading to the pileus's deformation. The PR-SOLOv2 instance segmentation model can more accurately separate them from the complex situations of adhesion between mushrooms, inclination, soil shielding on the surface, and uneven light, and obtain their masks. However, for mushrooms that are occluded or deformed due to aggregation growth, the mask obtained according to the image is only part of the shape of the mushroom. If the mask center point is directly extracted from the mushroom center point, there will be a significant error.

Therefore, given the above non-uniform and irregular mushroom growth shapes, especially the severe deformation and shape loss of densely overlapping mushroom contours, to better restore the actual mushroom contour and improve the center point positioning accuracy, this paper proposes a mushroom contour classification reconstruction method based on PR-SOLOv2 segmentation mask contour curvature and length, as shown in Figure 9.

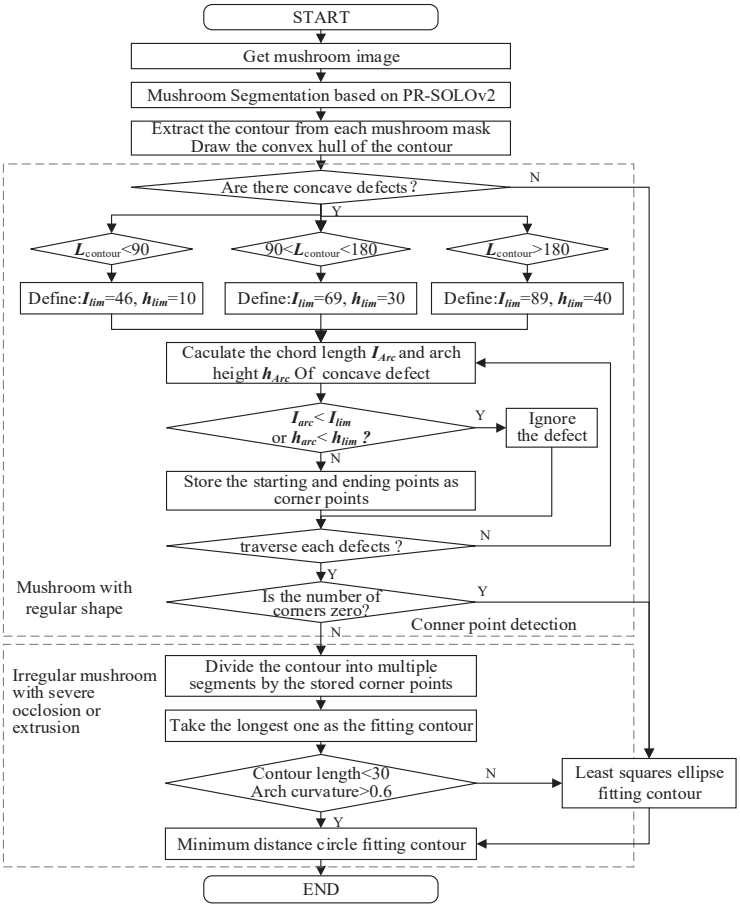


Figure 9. Flow chart of high-precision contour reconstruction based on PR-SOLOv2.

In this method, firstly, the PR-SOLOv2 algorithm is used to segment mushrooms and obtain their masks. Then, the mushroom mask contour based on the mask is extracted. Next, mushrooms are classified into regular and irregular shapes according to the concave defect length of the extracted contour. A mushroom with a regular shape is fitted with the least square ellipse. In contrast, a mushroom with an irregular shape and difficult contour fitting, seriously occluded or squeezed, is reconstructed with the longest contour extraction and classification reconstruction method based on corner segmentation.

5.1. Reconstruction of Regular Contours

The contour is considered regular with no or only shallow concave defects.

Firstly, the convex hull method is used to find out if there are any concave defects in the contour. Contours without concave defects are considered regular contours. In addition, although some mushrooms have concave defects, the chord length (I_{Arc}) and arch height (h_{Arc}) of the concave contour are relatively small compared to the entire mushroom length ($_rL_{contour}$), so such concave defects can be ignored. Mushroom contours that meet the above conditions can still be considered regular contours. Due to the varying sizes of mushrooms, to better determine the impact of concave contours on the fitting of mushrooms of different sizes, mushrooms are divided into three types based on the contour length ($L_{contour}$): small, medium, and large. Different boundary values are used to determine mushrooms of

different sizes. A mushroom with a contour length in the range [0,90) is defined as a small mushroom, [90,180) as a medium mushroom, and above 180 as a large mushroom. When I_{Arc} is less than 46 or h_{Arc} is less than 10 for a small mushroom, a concave contour can be ignored; when I_{Arc} is less than 69 or h_{Arc} is less than 30 for a medium mushroom, a concave contour can be ignored; when I_{Arc} is less than 89 or h_{Arc} is less than 40 for a large mushroom, a concave contour can be ignored. After the above judgment, if the concavities of the mushroom contour can be ignored, then this mushroom is defined as having a regular mushroom contour. Otherwise, it has an irregular contour.

For mushrooms with regular contours, they can be fitted by the least-squares ellipse method: Take the n contour points above it as $P_i(x_i, y_i)(i = 1, 2, \ldots, n)$, perform fitting operation according to the least-squares ellipse fitting objective function, calculate the ellipse position parameters and shape parameters, and reconstruct the contour of the bisporus mushroom with a smooth contour and regular shape.

5.2. Reconstruction of Irregular Contours

After segmenting the irregular contour of the overlapping mushroom, the obtained contour contains the contour fragment of the overlapping part of the other mushrooms covered. Suppose it is not processed and allowed to participate in the fitting together with the effective contour, such as the least-squares ellipse fitting. In that case, the fitted contour will differ from the actual contour. And the overlapped part cannot be well restored to the actual contour.

Given the above situation, this paper proposes the longest contour extraction and classification reconstruction approach relying on corner segmentation. The specific process is as follows:

(1) Detect corner points at abrupt changes in contour shape. If the concave contours cannot be ignored, define its start and end point as corner points;
(2) By corner coordinates, the coordinate line between adjacent corner coordinates is regarded as a sub-contour segment, and the whole mushroom contour is divided into N sub-contour segments;
(3) Calculate the length of each sub-contour segment and choose the longest sub-contour segment;
(4) Based on the longest sub-contour, the mushroom contour is classified and fitted according to the difference in arch curvature and length of the longest sub-contour.
(5) Calculate the arch curvature C of the contour (Equation (4)).

$$C = \frac{1}{R} \tag{4}$$

$$R = 0.5 * \left(\frac{z^2}{h_{Arc}} + h_{Arc} \right) \tag{5}$$

$$z = \sqrt{(x_n - x_1)^2 + (y_n - y_1)^2} \tag{6}$$

$$h_{Arc} = \max_i \left(\frac{S_i * 2}{z} \right) \tag{7}$$

$$S_i = \sqrt{p_i(p_i - z)(p_i - v_i)(p_i - m_i)}, i = 1, \ldots n \tag{8}$$

where R is the arcuate curvature radius of the longest sub-contour, v_i and m_i are, respectively, the length of both sides of the inscribed triangle formed by any point i (i = 1,2,. . .,n) on the longest sub-contour and its two upper endpoints, I_{Arc} is the chord length of the longest sub-contour and is also the bottom length of all inscribed triangles in the longest sub-contour, p_i is the half circumference of the inscribed triangle, S_i is the area of the

inscribed triangle, h_{Arc} is the height of the longest arch in the inscribed triangle on the longest sub-contour.

For the longest sub-contour with an arch curvature exceeding 0.6 and a length surpassing 30, the least-squares ellipse is employed to fit the contour. However, when dealing with the longest sub-contour having an arch curvature less than 0.6 or a length below 100, using the least-squares ellipse fitting method would result in a significant deviation of the fitting contour and its center point from the actual values. Consequently, the minimum-distance circle-fitting method is applied to reconstruct the mushroom contour in these two scenarios. N points (x_j, y_j) on the contour are selected, and the circle parameters are determined based on the sum of the absolute values of the distances between the data points and the circle, as depicted in Equation (9):

$$f = \sum \left| \sqrt{(x_j - x_c)^2 + (y_j - y_c)^2} - r \right| \tag{9}$$

where x_c and y_c are the center point of the fitted circle, r is the radius of the fitted circle, f is the minimum value of x_c, y_c and r are the best-fit parameters. For irregular contours, the longest sub-contour is extracted based on the corner segmentation proposed above, and then the target contours of overlapping and irregular mushrooms are reconstructed separately by least-squares ellipse fitting and minimum-distance circle-fitting according to the different situations of the bow curvature and length of the longest sub-contour.

6. Experimental Results and Discussion

6.1. Segmentation Experiment and Discussion Based on PR-SOLOv2

6.1.1. Model Training

To validate the results and advantages of the enhanced algorithm, this article uses Mask RCNN, YOLACT, SOLOv2, and PR-SOLOv2 image segmentation networks for training based on the mushroom datasets mentioned above to compare and verify their segmentation effects.

The model training environment was built under the Windows 10 operating system and a 1×GeForceRTX3090 NVIDIA graphics card from the United States, and the running memory was 24 GB. Framework Detectron2 of Facebook under the running architecture CUDA10.1 and CUDNN7.6.5 acceleration libraries of Python3.7, PyTorch1.6, and GPU were configured as the running framework of the model.

Training parameters were configured with a training batch size of 4. The neural network optimization algorithm employed stochastic gradient descent (SGD) to progress in the direction of the maximum slope at the current location. The initial learning rate was set at 0.01 and the momentum size at 0.9, and 5000 iterations of iterative training were carried out. In addition, the mesh size S of SOLOv2 and PR-SOLOv2 algorithms was (12,16,24,36,40).

6.1.2. Model Evaluation Metrics

The instance segmentation algorithm is essentially a problem of pixel classification and category recognition. Its evaluation metrics include Average Precision (AP), Intersection over Union (IOU), forecast time, and other common metrics.

AP represents the area value of the curve enclosed by the coordinate axis, and the precision–recall (PR) curve is plotted with the recall rate as the horizontal coordinate and the accuracy rate as the vertical coordinate. IOU denotes the intersection area of the target mask predicted by the model and the actual annotation box, along with the ratio of their combination. A preset judgment threshold is utilized, and when the IOU value exceeds this threshold, the model's prediction for the target is deemed relatively accurate.

The mushroom segmentation task necessitates predicting categories at the pixel level. Therefore, evaluation indicators such as the IOU threshold and the corresponding AP value are adopted. These metrics help distinguish the intersection ratio between the predicted segment and the ground truth. The average accuracies AP_{50} and AP_{75}, obtained using the

commonly used decision thresholds of 0.5 and 0.75 in the example segmentation task, are used to evaluate the segmentation accuracy of the model under different confidence levels.

6.1.3. Instance Segmentation Results and Discussion

(1) Experiment and result of dense overlapping mushroom segmentation

For the mushrooms that grow densely and form mutual adhesion or even mutual extrusion, the segmentation effect of different instance segmentation methods is shown in Figure 10. The segmentation result of Mask RCNN shows inaccurate segmentation of mushroom masks and missing identification of some mushrooms, as shown in Figure 10b. No. 1, 2, 3, and 4 mushrooms gather and grow, stick together, squeeze, and deform each other. There is no height difference between mushrooms, which leads to their indistinct boundary. No. 6 and No. 8 mushrooms were not identified because they were severely obscured by their neighbors, with only a small amount of fungus cap exposed and uneven light exposure. Due to the limited growth space of No. 5 and No. 7 mushrooms, the fruiting body can only grow slightly inclined, exposing the stalk. It is difficult to distinguish between the mushroom cap to be identified and the stem. Mask RCNN cannot handle these complex instance segmentation cases.

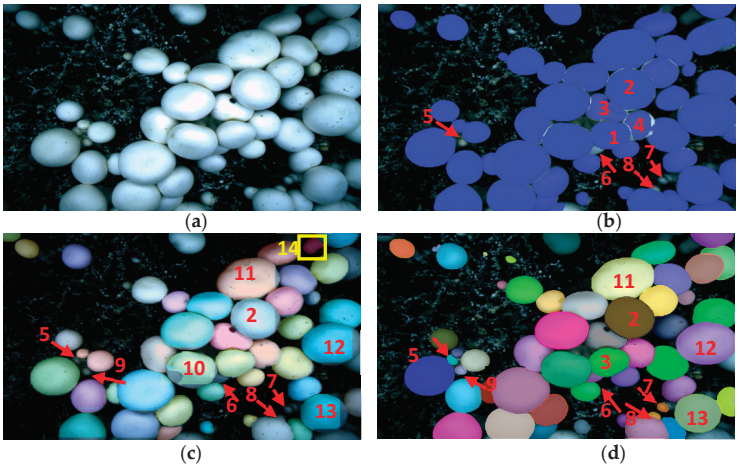


Figure 10. Segmentation results of densely overlapping mushrooms: (**a**) Original; (**b**) Mask RCNN; (**c**) YOLACT; (**d**) PR-SOLOv2.

The segmentation effect of the YOLACT algorithm is shown in Figure 10c. There were some cases in which the stalk was misidentified as a mushroom, some mushrooms were not identified, and the shape of the generated mushroom mask was very different from the actual mushroom shape. The stalk of mushroom No. 14, marked in yellow box in the picture, is misidentified as a mushroom; No. 5, 7, and 9 mushrooms that grow slightly at an angle and expose their stalks are missed; No. 6 and No. 8 mushrooms were seriously blocked by the large mushrooms growing around them, resulting in a small imaging area and uneven illumination and not being recognized. The shape of the segmentation mask generated by No. 2, 10, 11, 12, and 13 mushrooms is rectangular, greatly different from the actual mushroom shape. Like Mask RCNN, the YOLACT algorithm cannot solve the complex and challenging high-precision segmentation problems mentioned above.

In contrast is the PR-SOLOV2 algorithm for dense adhesion and extrusion deformation. It can be accurately segmented in complex segmentation situations, such as non-obvious features of the fungus cap and stalk, as shown in Figure 10d. The PR-SOLOv2 can accurately segment No. 1, 2, 3, and 4 mushrooms that grow densely and extrude each other without blurred edges. It can also separate No. 6 and No. 8 mushrooms that can be heavily obscured

and not easily identified. In addition, an excellent segmentation effect was also shown on No. 5, 7, and 9 mushrooms with no obvious characteristics of fungus cap and stalk, and no fungus stalk was misidentified as fungus cap. The mask shapes generated on No. 10, 11, 12 and 14 mushrooms will not appear rectangular like the YOLACT but will generate masks that conform to the actual shape of the mushrooms.

(2) Segmentation of a tilting mushroom

For the tilted mushroom, the segmentation problems of the Mask RCNN algorithm and the YOLACT algorithm are similar, as shown in Figure 11b,c. No. 1 and 2 mushrooms grow at a full tilt to the soil surface, exposing most of the stalks. The stalks look like tiny mushrooms. The growth posture of No. 3, 4, and 5 mushrooms was slightly inclined, revealing a small amount of stem. With inaccurate segmentation results, the mask RCNN and YOLACT algorithms also divided part of the stalk into the cap. No. 6 mushroom grows together with other mushrooms and hides on the stem of another mushroom. When the Mask RCNN and YOLACT algorithm segment the mushroom, in this case, part of the stalk is also divided into the fungus cap to obtain an unclear mask. The PR-SOLOv2 as shown in Figure 11d can correctly segment the stem with similar characteristics to the mushroom cap due to oblique growth, which increases the difficulty of identifying the mushroom. Including the correct segmentation of No. 1 and No. 2 of this oblique growthwhich is difficult to accurately identify. It can also accurately divide mushrooms No. 3, 4, 5, and 6, whose stalk and cap characteristics are difficult to distinguish, and get their precise edge.

(3) Segmentation result of tiny mushrooms

The identification effect of small mushrooms is shown in Figure 12. Mask RCNN and YOLACT have similar products in small mushroom recognition, as shown in Figure 12b,c. No. 1, 3, 4, 7, 8, 9, 11, 12, 13, 16, 17, and 19 mushrooms are not exposed to uniform light, the individual mushroom is small, and the characteristics of the fungus cap are not prominent. Therefore, it is easy to be mistaken for mycelium and miss recognition. Individual mushrooms No. 2, 5, 6, 10, 14, 15, 18, 20, and 21 are slightly larger, belonging to the small mushroom and large mushroom adhesion with a certain height difference. It is easy to identify the small mushroom in the stalk of the large mushroom, which also missed identification. However, the effect of the PR-SOLOv2 algorithm is shown in Figure 12d; it can correctly segment the mushrooms with smaller individuals and fewer features, whose characteristics are similar to mycelia. In addition, it can also accurately segment small mushrooms that are adhered to large mushrooms and have different heights and similar features.

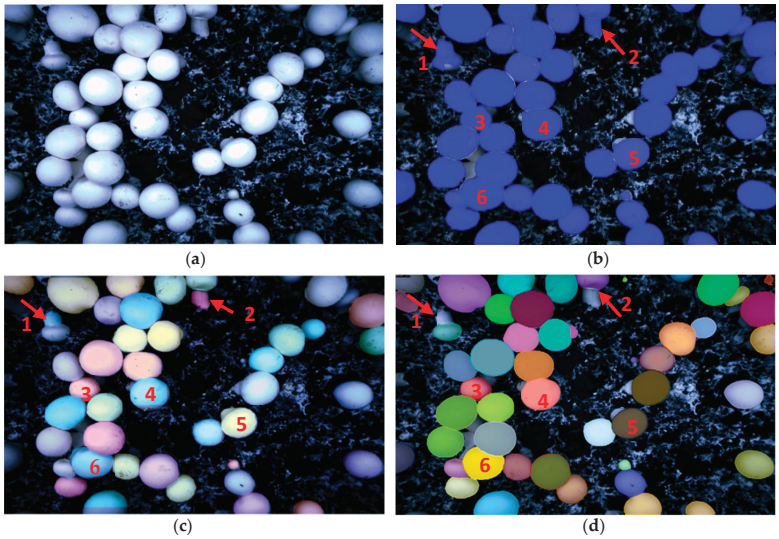


Figure 11. Segmentation result of tilting mushrooms: (**a**) Original; (**b**) Mask RCNN; (**c**) YOLACT; (**d**) PR-SOLOv2.

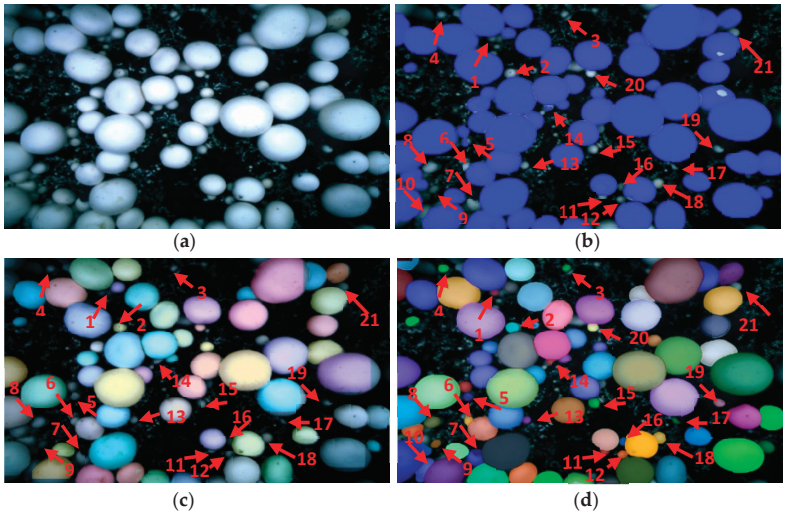


Figure 12. Segmentation result of tiny mushrooms: (**a**) Original; (**b**) Mask RCNN; (**c**) YOLACT; (**d**) PR-SOLOv2.

(4) Comparison of segmentation effects

The effect of instance segmentation on mushroom segmentation is tabulated in Table 2. The average segmentation accuracy of the Mask RCNN on the dense mushroom image segmentation task is 83.510%. The AP of YOLACT is 80.743%, while that of SOLOv2 is more than 90% (90.279%). The PR-SOLOv2 proposed in this paper has the highest AP among the four segmentation methods in dense mushroom segmentation, reaching 93.037%, which is about 10% higher than Mask RCNN and YOLACT, and about 3% higher than that of SOLOv2. When the IOU threshold is set to 50%, the AP performance of PR-SOLOv2 in the mushroom image segmentation task is as high as 99.056%. In addition,

the average segmentation time is 0.39 s, which is still better than the Mask RCNN and YOLACT algorithms.

Table 2. Effect of instance segmentation on mushroom segmentation.

Algorithm	P	R	AP (%)	AP_{50} (%)	AP_{75} (%)	Mean Time(s)
Mask RCNN	0.933	0.852	83.510	95.061	89.006	0.68
YOLACT	0.901	0.838	80.743	93.451	87.853	0.41
SOLOv2	0.959	0.917	90.279	96.103	92.747	0.36
PR-SOLOv2	0.982	0.945	93.037	99.056	95.249	0.39

(5) Comparison of segmentation effect under different light sources

In order to compare the segmentation effect of the model under the three light sources, this paper selects 60 pictures for experiment; the pictures are all from the camera used in the experiment; each picture has the same resolution and is divided into three groups according to the light sources. The overall results of the experiment are as follows: 651 examples of top light source images, 638 successful segmentations, segmentation success rate of 97.81%; 794 examples of side light source images, 772 successful segmentations, segmentation success rate of 97.23%; 1102 examples of dark light source images, 1088 successful segmentations, segmentation success rate of 98.73%, as shown in Figure 13.

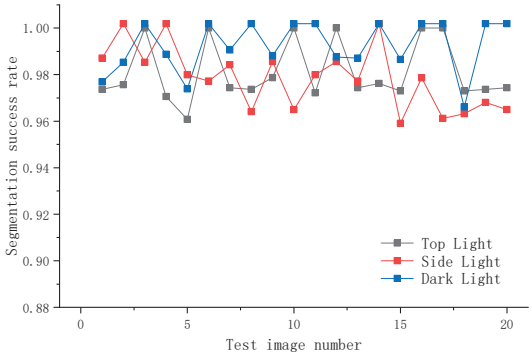


Figure 13. Segmentation results under different light sources.

The result shows that the segmentation accuracy of the side light source is the lowest, lower than that of the top light source, and lower than that of the dark light source. This paper concludes that since the side light source produces shadows on the mushroom surface when irradiating the mushroom, the light and dark demarcation line will affect the extraction of the mushroom edge features to a certain extent, so it will affect the final segmentation success rate. Based on the results of this experiment, considering that in the natural state, most of the stacked mushroom racks are illuminated by lights with side light sources, the direct photographs will affect the subsequent recognition and contour reconstruction. Therefore, in this paper, the camera is equipped with a parallel light source to assist illumination when taking pictures, to eliminate the influence of the side light source on the subsequent recognition and segmentation of mushrooms, and to improve the accuracy of recognition and contour reconstruction.

(6) Further validation of PR-SOLOv2 algorithm

There are 260 mushroom sample images of different mushroom rooms and layers collected in the actual production environment from two plants, totaling 9362 mushrooms. The mask RCNN, YOLACT, SOLOV2, and PR-SOLOv2 algorithms were used to segment the sample images. The experimental outcomes are shown in Table 3. It is observed that the

proposed approach achieves a segmentation success rate of 98.13%, which is significantly higher than those of the other existing works.

Table 3. Mushroom segmentation success rate of different algorithms.

Plant	Algorithm	Number	Num of Successful Segment Mushroom	Segmentation Success Rate (%)
Plant 1	Mask RCNN	4865	4209	86.52
	YOLACT	4865	3908	80.33
	SOLOv2	4865	4407	90.59
	PR-SOLOv2	4865	4784	98.34
Plant 2	Mask RCNN	4497	3853	85.68
	YOLACT	4497	3617	80.43
	SOLOv2	4497	4077	90.66
	PR-SOLOv2	4497	4403	97.91

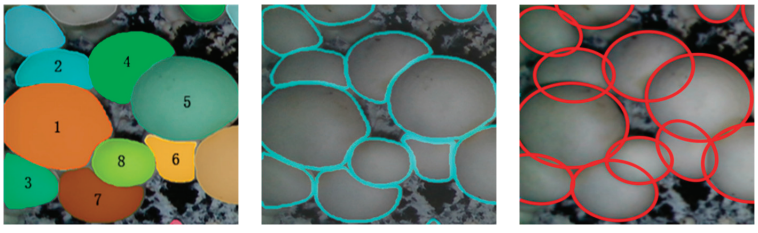
There are two main reasons for the unsuccessful segmentation: first, multiple large mushrooms are blocked at the same time, resulting in too serious small mushrooms; Second, a small mushroom with a large area of dark lines on the surface. In the marking of the dataset, the two cases are not common, resulting in poor learning effect of the segmentation features of the model for these two cases. In the future, the segmentation efficiency of the model can be improved by adding the marking of special cases.

6.2. Experimental Results and Discussion of Contour Reconstruction

The mask contour obtained by segmentation mask based on the PR-SOLOv2 algorithm proposed in this paper is adopted. The process and effect of the classification and reconstruction method according to the curvature and arc length of the mask contour are shown in Figure 14: (a) the fruit-body mask obtained by PR-SOLOv2 segmentation network processing proposed in this paper; (b) the extracted mask contour. Among them, No. 1, 5, and 8 are mushrooms whose average curvatures are 0.01317, 0.01367, and 0.02313, respectively, and corresponding lengths are 444, 427, and 246 pixels, respectively. They are all mushrooms whose average curvature is less than 0.103 and whose contour length is more than 100. No. 2, 3, 4, 6 and 7 are partially occluded fruiting bodies, whose average curvatures are 0.02378, 0.01787, 0.01648, 0.02087, and 0.02252, respectively, and corresponding contour lengths are 291, 257,335, 239,and 316 pixels, respectively, belonging to non-smooth occluded contours. Among them, the arch curvature of the largest contour fragment of No. 2, No.4, and No. 7 mushrooms is more significant than 0.9. The least-squares ellipse fitting was used. The No. 3 and No. 6 mushrooms were extruded to produce a large deformation, and their contour shape was irregular. However, the arch curvature of their longest contour was greater than 0.9, so the ellipse fitting was also adopted. The figure illustrates that both the regular mushroom and the severely blocked mushroom exhibit a well-matched contour when employing the contour reconstruction method proposed in this paper. The reconstruction effect is satisfactory.

More reconstruction effects of the mushroom cluster neutron entity are shown in Figure 15. In general, the contour reconstruction method in this paper can fit the contour of the densely overlapped fruit body well. For example, the surface damage of No. 1, 7, 8, 12, and 26 mushrooms and the surface dirt of No. 3 mushrooms will affect the image segmentation effect and lead to a poor subsequent reconstruction effect. The SOLOv2-PR algorithm proposed in this paper can accurately segment the mushroom mask and extract the edge of the mask. On this basis, the mushroom contour reconstruction can have a good effect. No. 6, 10, 13, 17, 18, 19, 20, 21, 22, 26, 23, 27, 28, and 29 mushrooms are mainly adhered to a mushroom around them or are partially blocked by a single mushroom next to them. Although these mushrooms are blocked or squeezed, the deformation of the mushrooms is not severe, and the extracted mushroom outline is roughly smooth and

regular. The contour reconstruction method proposed in this paper can achieve a better effect than mushroom contour reconstruction. In addition, for the mushrooms squeezed by two surrounding mushrooms, such as No. 4, 5, 9, 12, 15, 14, 16, 25, 31, 32, 34, and 33, the contour edge deformation is large, and the shape is irregular, so it is not easy to reconstruct the contour. The mushroom contour reconstruction method realized in this paper can restore the mushroom's true contour.



(**a**) (**b**) (**c**)

Figure 14. Reconstruction effect of overlapping extruded mushrooms: (**a**) PR-SOLOv2 mask; (**b**) Extract mask edges; (**c**) Contour reconstruction.

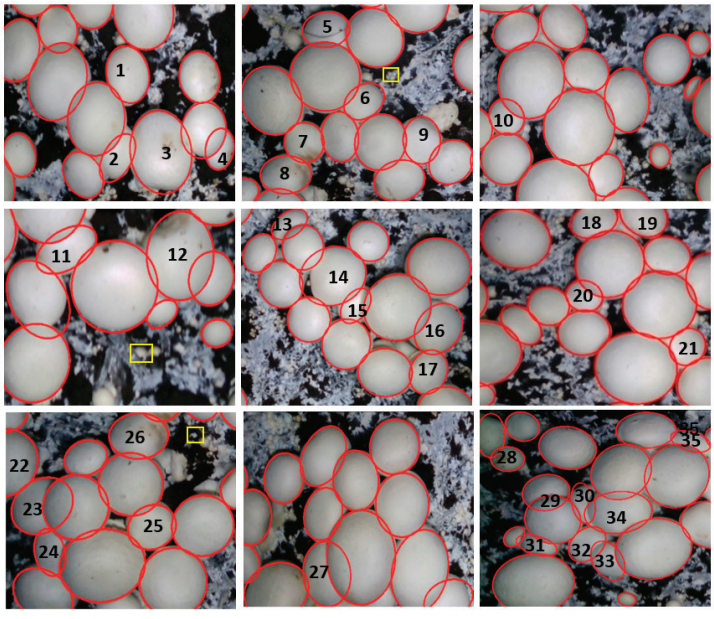


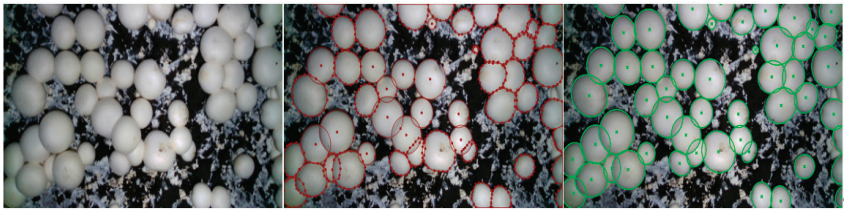
Figure 15. Mushroom contour reconstruction effect.

However, in cases of extremely dense growth, a majority of the fruiting bodies, like mushrooms numbered 2, 11, 15, 24, 30, and 35, are either blocked or severely deformed by the presence of multiple surrounding mushrooms. The PR-SOLOv2 algorithm is employed to segment the mask contour of these severely deformed mushrooms and subsequently reconstruct their contours. There may be some deviation between the fitted contour and the actual contour. However, there will be missed recognition for the mushroom fruiting bodies smaller than 30 pixels, such as the mushrooms in the area marked by the yellow rectangle box. However, these are tiny mushrooms with fresh buds, so the missing identification does not affect the picking or monitoring.

6.3. Verification of Center Coordinates

6.3.1. Center Point Ground Truth Determination Method

In order to find the ground true of center point of each fruiting body in the mushroom image, we use manual marking to determine the center point of each fruiting body, as shown in the Figure 16. The center point of the daughter body with regular contour is directly obtained by using circle annotation in labelme, and the daughter body with irregular contour is obtained by using polygon annotation in labelme, as close as possible to its true contour, as shown in Figure 16b. After all mushrooms are marked, the center point of the child body marked with a circle is directly output as its benchmark, and the center point of the child body marked with a polygon is obtained by ellipse fitting output, as shown in Figure 16c. The coordinates of all the central green 'x' in the final output image are used as the reference point of each sub-entity center.

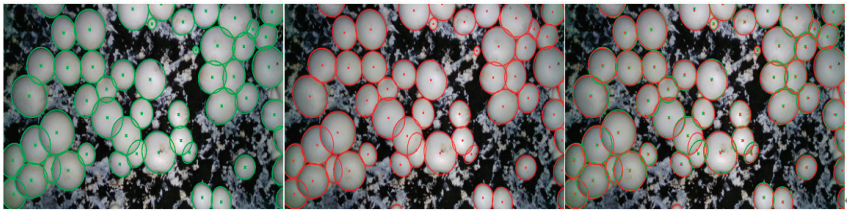


(**a**) (**b**) (**c**)

Figure 16. Benchmarking process for determining mushroom centers: (**a**) Original image; (**b**) Marking image; (**c**) Fitting contour map.

6.3.2. Calculation of Fitting Accuracy

In order to quantify the contour reconstruction accuracy, the center point positioning accuracy of the contour is used to characterize it. Firstly, the deviation from the center datum is compared from the actual fitting effect graph, as shown in Figure 17: the center point of the actual contour using manual labeling, as shown in Figure 17a; and the center point of the contour reconstructed by PR-SOLOv2, as shown in Figure 17b; for comparison, the comparison effect is shown in Figure 17c. It can be seen that the contour reconstructed by PR-SOLOv2 fits better with the actual contour manually labeled, and the center point deviation is very small.



(**a**) (**b**) (**c**)

Figure 17. Comparative effect of fitting: (**a**) Manually labeled contours; (**b**) Reconstructed contours by this paper's method; (**c**) Comparison result.

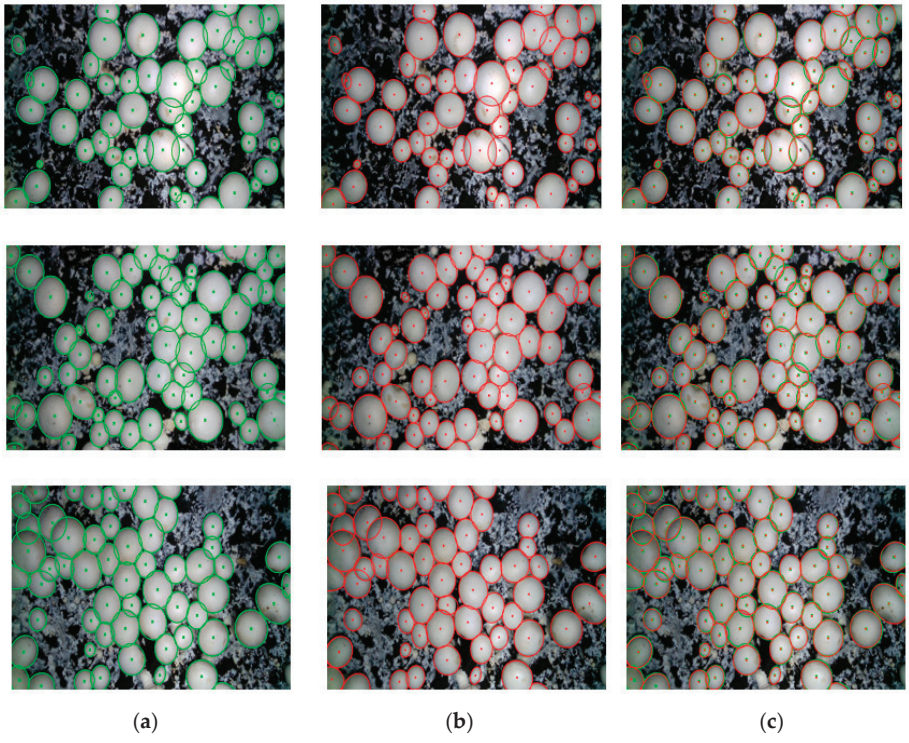
In order to calculate the deviation more quantitatively, the center point positioning accuracy is calculated using the 2D coordinate deviation rate (CDR) [24] as depicted in Equation (10).

$$E = \left(\left| \frac{c_j - c_i}{W} \right| + \left| \frac{r_j - r_i}{H} \right| \right) \times 100\% \tag{10}$$

where r_j and c_j are the row and column coordinates of the center point of the mushroom contour fitted by manual marking, r_i and c_i are the row and column coordinates of the

center point of the mushroom obtained by the research method in this paper, and W and H are the width and height of the mushroom image, respectively.

The contour reconstruction comparison results of three more dense onsite mushroom images (169 mushrooms in total) are provided to further verify the effectiveness of the contour reconstruction method in this paper, as shown in Figure 18. The corresponding CDR comparison of the four images is shown in Figure 19.



(**a**) (**b**) (**c**)

Figure 18. Comparison effect of contour reconstruction by this paper's method and manually labeled contours: (**a**) Manually labeled contours; (**b**) Reconstructed contours by this paper's method; (**c**) Comparison result.

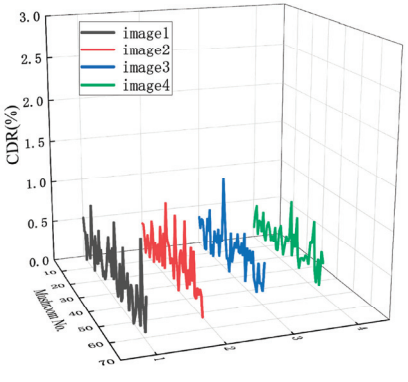


Figure 19. CDR plots of the four pictures.

As shown in Table 4, the average CDR between the centroid of the contour fitted by the method in this paper and the actual contour is 0.30%, which shows that the centroid positioning accuracy is higher, with better results.

Table 4. Mean deviation statistics for the average CDR coordinates of the four pictures.

Image Number	1	2	3	4	Average
The contour number	63	54	52	50	
Average CDR	0.28%	0.32%	0.31%	0.30%	0.3%

7. Conclusions

This paper presents a high-precision segmentation and contour reconstruction method for densely overlapping mushrooms based on an improved SOLOv2 model (PR-SOLOv2), aiming to improve the segmentation accuracy and center point positioning accuracy of highly dense mushrooms. The experimental results show significant improvement. The PR-SOLOv2 model improves segmentation accuracy to 93.04% and segmentation rate by 98.13%, about 10% higher than Mask RCNN and YOLACT. Meanwhile, the center positioning error is as low as 0.3%, which proved that the proposed contour reconstruction method in this study could accurately segment dense, overlapped, extruded, and tilting fruiting bodies, effectively reconstruct mushroom contours that are severely occluded and compressed, and achieve accurate center point positions. It is not only suitable for the high-precision identification and positioning of the fruiting bodies of various edible fungi but also those of other densely overlapping spherical-like fruits. However, there are still some limitations. For example, for mushrooms with too a high tilt angle, occasionally it will be mistaken to identify the stalk of the mushroom as a fungus cap. When the fruiting body of the mushroom is severely obstructed, the reconstructed mushroom contour may have some relatively larger deviations from the actual mushroom contour, which needs to be further improved. In the future, we will continue to improve the structure of the segmentation model, and even try some new occlusion segmentation algorithms such as amodal instance segmentation, to further enhance the segmentation effect and reduce processing time, and optimize the reconstruction fitness of the contour reconstruction algorithm for contours covered by multiple positions or severely obstructed, to further enhance center point position accuracy.

Author Contributions: S.Y.: Writing—original draft, writing—review and editing, conceptualization, methodology, investigation, data curation, formal analysis, software, validation, production of related equipment, funding acquisition. J.Z.: writing—original draft, writing—review and editing, data curation, validation, software. J.Y.: investigation, writing-review and editing, conceptualization. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by the Shanghai Agriculture Science and Technology Innovation Program, Grant No. I2023003; Shanghai Science and Technology Innovation Action Plan-Agriculture, Grant No. 21N21900600; Major Scientific and Technological Innovation Project of Shandong Province, Grant No. 2022CXGC010609.

Institutional Review Board Statement: Not applicable.

Data Availability Statement: Anyone can access our data by sending an email to szyang@sspu.edu.cn.

Acknowledgments: We acknowledge the Shanghai Lianzhong edible fungus professional cooperative for the use of their mushrooms and facilities.

Conflicts of Interest: The authors declare no conflicts of interest.

References

1. Diego, C.Z.; Pardo-Giménez, A. Edib SOLOv2: Dynamic and fast instance segmentation le and medicinal mushrooms (technology and applications); Current overview of mushroom production in the world. In *Edible and Medicinal Mushrooms: Technology and Applications*; Wiley: Hoboken, NJ, USA, 2017; pp. 5–13.
2. Zhang, J.; Cai, W.; Huang, C. *Chinese Edible Fungus Cultivation*, 1st ed.; China Agriculture Press: Beijing, China, 2020; Chapter 1.
3. Reed, J.N.; Miles, S.J.; Butler, J.; Baldwin, M.; Noble, R. Automatic mushroom harvester development. *J. Agric. Eng. Res.* **2001**, *78*, 15–23. [CrossRef]
4. Rowley, J.H. Developing Flexible Automation for Mushroom Harvesting (*Agaricus bisporus*): Innovation Report. Ph.D. Thesis, University of Warwick, Warwick, UK, 2009.
5. Yang, C.H.; Xiong, L.Y.; Wang, Z.; Wang, Y.; Shi, G.; Kuremot, T.; Yang, Y. Integrated detection of citrus fruits and branches using a convolutional neural network. *Comput. Electron. Agric.* **2020**, *174*, 105469. [CrossRef]
6. Yang, S.; Jia, B.; Yu, T.; Yuan, J. Research on multiobjective optimization algorithm for cooperative harvesting trajectory optimization of an intelligent multiarm straw-rotting fungus harvesting robot. *J. Agric.* **2022**, *12*, 986. [CrossRef]
7. Noble, R.; Reed, J.N.; Miles, S.; Jackson, A.F.; Butler, J. Influence of mushroom strains and population density on the performance of a robotic harvester. *J. Agric. Eng. Res.* **1997**, *68*, 215–222. [CrossRef]
8. Slaughter, D.C.; Harrell, R.C. Color vision in robotic fruit harvesting. *Trans. ASAE* **1987**, *30*, 1144–1148. [CrossRef]
9. Gené-Mola, J.; Gregorio, E.; Guevara, J.; Auat, F.; Sanz-Cortiella, R.; Escolà, A.; Rosell-Polo, J.R. Fruit detection in an apple orchard using a mobile terrestrial laser scanner. *Biosyst. Eng.* **2019**, *187*, 171–184. [CrossRef]
10. Lin, G.; Tang, Y.; Zou, X.; Li, J.; Xiong, J. In-field citrus detection and localisation based on RGB-D image analysis. *Biosyst. Eng.* **2019**, *186*, 34–44. [CrossRef]
11. Magalhães, S.A.; Moreira, A.P.; Santos, F.N.D.; Dias, J. Active Perception Fruit Harvesting Robots—A Systematic Review. *J. Intell. Robot. Syst.* **2022**, *105*, 1–22. [CrossRef]
12. Montoya-Cavero, L.E.; de León Torres, R.D.; Gómez-Espinosa, A.; Cabello, J.A.E. Vision systems for harvesting robots: Produce detection and localization. *Comput. Electron. Agric.* **2021**, *192*, 106562. [CrossRef]
13. Yan, B.; Fan, P.; Lei, X.; Liu, Z.; Yang, F. A real-time apple targets detection method for picking robot based on improved YOLOv5. *Remote Sens.* **2021**, *13*, 1619. [CrossRef]
14. Krizhevsky, A.; Sutskever, I.; Hinton, G.E. ImageNet classification with deep convolutional neural networks. *Commun. ACM* **2017**, *60*, 84–90. [CrossRef]
15. Onishi, Y.; Yoshida, T.; Kurita, H.; Fukao, T.; Arihara, H.; Iwai, A. An automated fruit harvesting robot by using deep learning. *ROBOMECH J.* **2019**, *6*, 1–8. [CrossRef]
16. Zhou, C.; Hu, J.; Xu, Z.; Yue, J.; Ye, H.; Yang, G. A novel greenhouse-based system for the detection and plumpness assessment of strawberry using an improved deep learning technique. *Front. Plant Sci.* **2020**, *11*, 559. [CrossRef] [PubMed]
17. Ganesh, P.; Volle, K.; Burks, T.F.; Mehta, S.S. Deep orange: Mask R-CNN based orange detection and segmentation. *IFAC PapersOnLine* **2019**, *52*, 70–75. [CrossRef]
18. Ghiani, L.; Sassu, A.; Palumbo, F.; Mercenaro, L.; Gambella, F. In-field automatic detection of grape bunches under a totally uncontrolled environment. *Sensors* **2021**, *21*, 3908. [CrossRef]
19. Yu, Y.; Zhang, K.; Yang, L.; Zhang, D. Fruit detection for strawberry harvesting robot in non-structural environment based on Mask-RCNN. *Comput. Electron. Agric.* **2019**, *163*, 104846. [CrossRef]
20. Zhang, K.; Lammers, K.; Chu, P.; Li, Z.; Lu, R. System design and control of an apple harvesting robot. *Mechatronics* **2021**, *79*, 102644. [CrossRef]
21. Tillett, R.D.; Batchelor, B.G. An algorithm for locating mushrooms in a growing bed. *Comput. Electron. Agric.* **1991**, *6*, 191–200. [CrossRef]
22. Yu, G.H.; Luo, J.M.; Zhao, Y. Region marking technique based on sequential scan and segmentation method of mushroom images. *Trans. Chin. Soc. Agric.* **2006**, *22*, 4.
23. Yang, Y.Q.; Ye, M.; Lu, Y.H.; Ren, S.G. Localization algorithm based on corner density detection for overlapping mushroom image. *Comput. Syst. Appl.* **2018**, *24*, 119–125.
24. Yang, S.; Ni, B.; Du, W.; Yu, T. Research on an improved segmentation recognition algorithm of overlapping *Agaricus bisporus*. *Sensors* **2022**, *22*, 3946. [CrossRef] [PubMed]
25. Yin, H.; Yi, W.; Hu, D. Computer vision and machine learning applied in the mushroom industry: A critical review. *Comput. Electron. Agric.* **2022**, *198*, 107015. [CrossRef]
26. Mukherjee, A.; Sarkar, T.; Chatterjee, K.; Lahiri, D.; Nag, M.; Rebezov, M.; Lorenzo, J.M. Development of artificial vision system for quality assessment of oyster mushrooms. *Food Anal. Methods* **2022**, *15*, 1663–1676. [CrossRef]
27. Wang, Z.J.; Zhang, L.L.; Li, S.F.; Zhang, Q.Y.; Fu, Q.Q.; Liu, S.J. Identification and classification of eatable fungi based on machine learning algorithm. *J. Fuyang Teach. Coll. Nat. Sci.* **2021**, *38*, 7.
28. Zhou, J.; Ding, W.; Zhu, X.; Cao, J.; Niu, X. Evaluation on formation rate of *Pleurotus eryngii* primordium under different humidity conditions by computer vision. *J. Zhejiang Univ. Sci. Agric. Life Sci.* **2017**, *43*, 262–272. [CrossRef]
29. Huang, B.R.; Chen, J.Y.; Lai, X.Y.; Chen, G.W.; Huang, K.W. Application of Deep Learning for Mushrooms Cultivation. In Proceedings of the 2021 International Conference on Technologies and Applications of Artificial Intelligence (TAAI), Taipei, Taiwan, 21 January 2021; pp. 250–253.

30. Huang, M.; Jiang, X.; He, L.; Choi, D.; Pecchia, J.; Li, Y. Development of a robotic harvesting mechanism for button mushrooms. *Trans. ASABE* **2017**, *64*, 565–575. [CrossRef]
31. Ketwongsa, W.; Boonlue, S.; Kokaew, U. A new deep learning model for the classification of poisonous and edible mushrooms based on improved AlexNet convolutional neural network. *J. Appl. Sci.* **2022**, *12*, 3409. [CrossRef]
32. Sevilla, W.H.; Hernandez, R.M.; Ligayo, M.A.D.; Costa, M.T.; Quismundo, A.Q. Machine vision recognition system of edible and poisonous mushrooms using a small training set-based deep transfer learning. In Proceedings of the 2022 International Conference on Decision Aid Sciences and Applications (DASA), Chiangrai, Thailand, 23–25 March 2022; pp. 1701–1705.
33. Zha, L.; Gong, Z.J.; Shao, S.; Li, Z.P.; Yu, C.X.; Yang, H.L. An anchor-free network for detection of *Volvariella volvacea* growth status. *Acta Edulis Fungi* **2022**, *29*, 8.
34. Lee, C.H.; Choi, D.; Pecchia, J.; He, L.; Heinemann, P. Development of a mushroom harvesting assistance system using computer vision. In Proceedings of the 2019 ASABE Annual International Meeting, American Society of Agricultural and Biological Engineers (ASABE), Boston, MA, USA, 7–10 July 2019; pp. 1–10.
35. Rong, J.; Wang, P.; Yang, Q.; Huang, F. A field-tested harvesting robot for oyster mushroom in greenhouse. *J. Agron.* **2021**, *11*, 1210. [CrossRef]
36. Zheng, Y.Y.; Zheng, L.X. Research on *Flammulina velutipes* head detection algorithm based on deep learning. *Mod. Comput.* **2020**, *30*, 23–29.
37. Lu, C.P.; Liaw, J.J. A novel image measurement algorithm for common mushroom caps based on convolutional neural network. *Comput. Electron. Agric.* **2020**, *171*, 105336. [CrossRef]
38. Masoudian, A. Computer Vision Algorithms for an Automated Harvester. Ph.D. Thesis, The University of Western Ontario, London, ON, Canada, 2013.
39. Sert, E.; Okumus, I.T. Segmentation of mushroom and cap width measurement using modified K-means clustering algorithm. *Adv. Electr. Electron. Eng.* **2014**, *12*, 354–360. [CrossRef]
40. Soomro, S.; Munir, A.; Choi, K.N. Hybrid two-stage active contour method with region and edge information for intensity inhomogeneous image segmentation. *PLoS ONE* **2018**, *13*, e0191827. [CrossRef] [PubMed]
41. Pchitskaya, E.; Bezprozvanny, I. Dendritic spines shape analysis—Classification or clusterization? Perspective. *Front. Synaptic Neurosci.* **2020**, *12*, 31. [CrossRef]
42. Ji, J.; Sun, J.; Zhao, K.; Jin, X.; Ma, H.; Zhu, X. Measuring the cap diameter of white button mushrooms (*Agaricus bisporus*) by using depth image processing. *Appl. Eng.* **2021**, *37*, 623–633. [CrossRef]
43. Arjun, A.D.; Chakraborty, S.K.; Mahanti, N.K.; Kotwaliwale, N. Non-destructive assessment of quality parameters of white button mushrooms (*Agaricus bisporus*) using image processing techniques. *J. Food Sci. Technol.* **2022**, *59*, 2047–2059. [CrossRef]
44. Baisa, N.L.; Al-Diri, B. Mushrooms Detection, Localization and 3D Pose Estimation using RGB-D Sensor for Robotic-picking Applications. *arXiv* **2022**, arXiv:2201.02837.
45. Zhao, K.; Zhang, M.; Ji, J.; Sun, J.; Ma, H. Whiteness measurement of *Agaricus bisporus* based on image processing and color calibration model. *Food Meas.* **2023**, *17*, 2152–2161. [CrossRef]
46. Retsinas, G.; Efthymiou, N.; Anagnostopoulou, D.; Maragos, P. Mushroom detection and three dimensional pose estimation from multi-view point clouds. *Sensors* **2023**, *23*, 3576. [CrossRef]
47. Zhang, J.; Li, C.; Rahaman, M.M.; Yao, Y.; Ma, P.; Zhang, J.; Zhao, X.; Jiang, T.; Grzegorzek, M. A comprehensive survey with quantitative comparison of image analysis methods for microorganism biovolume measurements. *Arch. Comput. Methods Eng.* **2023**, *30*, 639–673. [CrossRef]
48. Chen, C.; Wang, F.; Cai, Y.; Yi, S.; Zhang, B. An Improved YOLOv5s-Based *Agaricus bisporus* Detection Algorithm. *J. Agron.* **2023**, *13*, 1871. [CrossRef]
49. Chen, C.; Yi, S.; Mao, J.; Wang, F.; Zhang, B.; Du, F. A Novel Segmentation Recognition Algorithm of *Agaricus bisporus* Based on Morphology and Iterative Marker-Controlled Watershed Transform. *J. Agron.* **2023**, *13*, 347. [CrossRef]
50. Yin, H.; Xu, J.; Wang, Y.; Hu, D.; Yi, W. A Novel Method of Situ Measurement Algorithm for *Oudemansiella raphanipies* Caps Based on YOLO v4 and Distance Filtering. *Agronomy* **2022**, *13*, 134. [CrossRef]
51. Yin, H.; Yang, S.; Cheng, W.; Wei, Q.; Wang, Y.; Xu, Y. AC R-CNN: Pixelwise Instance Segmentation Model for *Agrocybe cylindracea* Cap. *Agronomy* **2023**, *14*, 77. [CrossRef]
52. Wang, T.; Liew, J.H.; Li, Y.; Chen, Y.; Feng, J. SODAR: Segmenting objects by dynamically aggregating neighboring mask representations. *arXiv* **2022**, arXiv:2202.07402.
53. Wang, X.; Kong, T.; Shen, C.; Jiang, Y.; Li, L. SOLO: Segmenting objects by locations. In Proceedings of the European Conference on Computer Vision (ECCV), Glasgow, UK, 23–28 August 2020; Springer: Cham, Switzerland, 2020; pp. 649–665.
54. Wang, X.; Zhang, R.; Kong, T.; Li, L.; Shen, C. SOLOv2: Dynamic and fast instance segmentation. *Adv. Neural Inf. Process. Syst.* **2020**, *33*, 17721–17732.
55. Long, J.; Shelhamer, E.; Darrell, T. Fully convolutional networks for semantic segmentation. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Boston, MA, USA, 12 June 2015; pp. 3431–3440.
56. Li, T.Y.; Jiang, W.W.; Xing, J.T.; Xu, Z.; Zhang, L.; Tian, F. Intrusion detection of offshore area perimeter in oilfield operation site. *Comput. Syst. Appl.* **2022**, *31*, 236–241.
57. Ji, J.T.; Liu, X.H.; Zhao, K.X. Automatic rumen filling scoring method for dairy cows based on SOLOv2 and cavity feature of point cloud. *Trans. Chin. Soc. Agric. Eng.* **2022**, *38*, 186–197.

58. Kirillov, A.; Wu, Y.; He, K.; Girshick, R. PointRend: Image segmentation as rendering. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), Seattle, WA, USA, 13 June 2020; pp. 9799–9808.
59. Lin, T.Y.; Goyal, P.; Girshick, R.; He, K.; Dollár, P. Focal loss for dense object detection. In Proceedings of the IEEE International Conference on Computer Vision (ICCV), Venice, Italy, 22 October 2017; pp. 2980–2988.

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.

agriculture

MDPI

Article

Model and Parameter Adaptive MPC Path Tracking Control Study of Rear-Wheel-Steering Agricultural Machinery

Meng Wang [1,2], Changhe Niu [3], Zifan Wang [1], Yongxin Jiang [3], Jianming Jian [1] and Xiuying Tang [1,*]

[1] College of Engineering, China Agricultural University, Beijing 100083, China; wm15@xjie.edu.cn (M.W.); s20213071234@cau.edu.cn (Z.W.); jamesjian@cau.edu.cn (J.J.)
[2] School of Mechanical and Electrical Engineering, Xinjiang Institute of Engineering, Urumqi 830023, China
[3] Research Institute of Agricultural Mechanization, Xinjiang Academy of Agricultural Sciences, Urumqi 830091, China; xjnch@xaas.ac.cn (C.N.); jyx@xaas.ac.cn (Y.J.)
* Correspondence: txying@cau.edu.cn; Tel.: +86-159-1060-4652

Abstract: To further enhance the precision and the adaptability of path tracking control, and considering that most of the research is focused on front-wheel steering, an adaptive parametric model predictive control (MPC) was proposed for rear-wheel-steering agricultural machinery. Firstly, the kinematic and dynamic models of rear-wheel-steering agricultural machinery were established. Secondly, the influence laws of curvature and velocity on the prediction horizon N_p, control horizon N_c, and preview value N_{pre} were obtained by simulating and analyzing the factors influencing the MPC tracking effect. The results revealed that raising N_{pre} can improve curve tracking performance. N_p was correlated negatively with the curvature change, whereas N_c and N_{pre} were positively connected. N_p, N_c, and N_{pre} were correlated positively with the velocity change. Then, the parameters for self-adaptation of N_p, N_c, and N_{pre} were accomplished via fuzzy control (FC), and particle swarm optimization (PSO) was utilized to optimize the three parameters to determine the optimal parameter combination. Finally, simulation and comparative analysis were conducted to assess the tracking effects of the manual tuning MPC, the FC_MPC, and the PSO_MPC under U-shaped and complex curve paths. The results indicated that there was no significant difference and all three methods achieved better tracking effects under no disturbance, with the mean absolute value of lateral error $\leq$0.18 cm, standard deviation $\leq$0.37 cm, maximum deviation of U-shaped path <2.38 cm, and maximum deviation of complex curve path <3.15 cm. The mean absolute value of heading error was $\leq$0.0096 rad, the standard deviation was $\leq$0.0091 rad, and the maximum deviation was <0.0325 rad, indicating that manual tuning can find optimal parameters, but with high uncertainty and low efficiency. However, FC_MPC and PSO_MPC have better adaptability and tracking performance compared to the manual tuning MPC with fixed horizons under variable-speed disturbance and are more able to meet the actual needs of agricultural machinery operations.

Keywords: path tracking; rear-wheel steering model; MPC; preview; fuzzy control; PSO

Citation: Wang, M.; Niu, C.; Wang, Z.; Jiang, Y.; Jian, J.; Tang, X. Model and Parameter Adaptive MPC Path Tracking Control Study of Rear-Wheel-Steering Agricultural Machinery. *Agriculture* **2024**, *14*, 823. https://doi.org/10.3390/agriculture14060823

Academic Editor: Valentin Vlăduț

Received: 9 April 2024
Revised: 16 May 2024
Accepted: 21 May 2024
Published: 24 May 2024



Copyright: © 2024 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (https://creativecommons.org/licenses/by/4.0/).

1. Introduction

The autonomous driving of agricultural machinery is critical to accomplish the "replacement of manual labor with machines" and answer the question of "who will farm in the future" [1]. Path tracking control is at the heart of autonomous driving technology, which is widely used in farming, planting, management, harvesting, transportation, and other aspects of agricultural production to improve work quality and efficiency and to reduce the waste of agricultural production materials [2]. Typical path tracking control methods include geometric model-based PP [3] and Stanley [4], etc., kinematic and dynamic model-based PID [5], LQR [6], SMC [7], and MPC [8], etc., as well as model-free based fuzzy control [9], reinforcement learning [10], and neural networks [11], etc.

MPC is the most widely used control method in practice other than PID, according to a survey by the International Federation of Automatic Control [12]. MPC can handle

constraint problems in real time and has low requirements for the accuracy of the control model. It is suitable for both linear and nonlinear systems, but tracking accuracy is difficult to guarantee at high speeds [13,14]. Therefore, it has unique advantages for agricultural machinery path tracking with relatively low operation speed, complex ground conditions, and difficulty in establishing accurate models. Xue et al. constructed a kind of MPC controller utilizing a linear time-varying model for orchard tractors, with average lateral errors of 7 cm and 13 cm for curve tracking at 2.5 m/s and 5 m/s, respectively [15]. He et al. built a kind of MPC controller for rice transplanters upon a kinematic model with attitude correction, and field experiments indicated that the average absolute error was 3.3 cm under a three-line linear path [16].

The initial parameter settings and weights of the components in the optimization objective have a significant impact on the computational results and directly affect the final tracking effect since MPC is essentially an optimization solution method. Neural networks, fuzzy control, and intelligent search algorithms are commonly used parameter adaptation methods. Different methods have their benefits as well as limitations. The selection process is primarily oriented to the actual demand (e.g., control accuracy, stability, adaptability, etc.), and the control results (e.g., tracking error) are measured as the indexes. Simultaneously, the computational capability of the implementation hardware is also taken into account. Zhou et al. applied the genetic algorithm to calculate the optimal time-domain parameters for articulated steering tractors, with an average lateral error of 2.1 cm under U-shaped path tracking, while the traditional MPC was 7.53 cm [17]. Wang et al. utilized the fuzzy principle to achieve the adaptive error weight coefficients of MPC, and the simulation results demonstrated the superiority of the upgraded MPC tracking effect over the PP controller [18]. Shi et al. used PSO-BPNN to improve MPC in the prediction and control horizon under different road adhesion coefficients [19]. Guan et al. dynamically adjusted the MPC's feedforward factor and lateral error weight based on BPNN. The maximum lateral error was 5.13 cm under double-lane path tracking and 9.23 cm for conventional MPC when the speed was 28 km/h [20]. Liang et al. used RBF neural networks to adaptively compensate for errors caused by changes in the curvature of complex paths at high speed, with a maximum error of 0.285 m [21]. Lin et al. adopted fuzzy control to optimize the prediction and control horizon online with v_x and v_y as inputs and the horizon factor as output [22]. Dai et al. proposed a kind of MPC with adaptive preview characteristics, and the maximum lateral error was 0.04 m at 10 m/s under double-lane path tracking by the simulation outcomes [23]. Choi et al. updated the sample time in the subsequent iteration based on the control inputs in each iteration of MPC and simulated the superiority of this method [24]. Liu et al. proposed a two-layer MPC with curvature adaption, which has strong adaptability at high speeds [25]. As can be seen from the above, traditional MPC with fixed parameters has poor adaptability to working conditions. The adaptation of key parameters such as time domain and error weights has become a research hotspot, but there is still significant space for improvement in terms of control precision and applicability, making this a crucial area for further study.

The steering methods of wheeled agricultural machinery mainly include front-wheel steering (such as tractors, transplanters, etc.), rear-wheel steering (such as grain harvesters, residual film recyclers, etc.), and four-wheel steering (such as high-clearance sprayers [26], etc.). Current research on agricultural vehicle models mainly focuses on front-wheel steering [27], with less research on the kinematic and dynamic models of rear-wheel-steering agricultural machinery, blurring or ignoring the difference between them, and equating them for control research.

To further improve the precision of path tracking and the adaptability of the control system to various operating conditions, and given that most research focuses on front-wheel steering, this study accomplished the following: (1) established the kinematic model and dynamic model of rear-wheel-steering agricultural machinery; (2) simulated and analyzed the key factors affecting the tracking effect, and discovered the interaction laws; (3) achieved the self-adaptation of the key parameters of MPC utilizing the fuzzy control

and PSO, and then designed the path-tracking controller to compare and analyze the parameter adaptive MPC's tracking effect with that of the traditional manual tuning MPC, proving the superiority of the designed controller.

2. Materials and Methods

2.1. Vehicle Model of Rear-Wheel-Steering Agricultural Machinery

2.1.1. Kinematic Model

Figure 1 represents the kinematic model of agricultural machinery schematically.

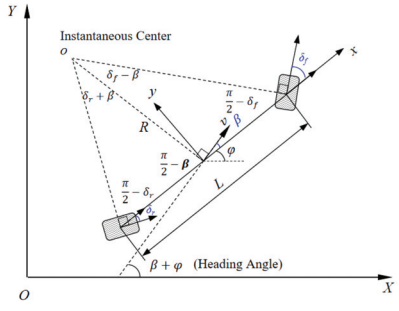


Figure 1. Schematic diagram of the kinematic model for agricultural machinery. Note: φ is the yaw angle; β is the center of mass lateral deviation angle; L is the wheelbase; v is the center of mass velocity; δ_f and δ_r are front- and rear-wheel angles, respectively; and R is the instantaneous turning radius.

The vehicle's kinematic equation is as follows, according to the geometric relationship of the steering:

$$\begin{cases} \dot{x} = v\cos(\beta + \phi) \\ \dot{y} = v\sin(\beta + \phi) \\ \dot{\phi} = \frac{v(\tan\delta_f + \tan\delta_r)\cos\beta}{L} \end{cases} \tag{1}$$

Under low-speed conditions (the operation speed of the agricultural machinery is usually <3 m/s), assuming that the vehicle has no side-slip, i.e., $v_y = 0$, then $\beta = \arctan\frac{v_y}{v_x} = 0$. In agricultural machinery with rear-wheel steering, the front-wheel steering angle $\delta_f = 0$; therefore, its kinematic equation is as follows:

$$\begin{cases} \dot{x} = v\cos\phi \\ \dot{y} = v\sin\phi \\ \dot{\phi} = \frac{v\tan\delta_r}{L} \end{cases} \tag{2}$$

Select $\chi = [x, y, \phi]^T$ as the state variables and $u = [v, \delta]^T$ as the control variables. The error state equation is obtained by performing a Taylor series expansion at the reference point, neglecting high-order terms, and calculating the Jacobian matrix.

$$\dot{\widetilde{\chi}} = \begin{bmatrix} 0 & 0 & -v_r\sin\phi_r \\ 0 & 0 & v_r\cos\phi_r \\ 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} x - x_r \\ y - y_r \\ \phi - \phi_r \end{bmatrix} + \begin{bmatrix} \cos\phi_r & 0 \\ \sin\phi_r & 0 \\ \frac{\tan\delta_r}{L} & \frac{v_r}{L\cos^2\delta_r} \end{bmatrix} \begin{bmatrix} v - v_r \\ \delta - \delta_r \end{bmatrix} = a\widetilde{\chi} + b\widetilde{u} \tag{3}$$

The error state equation is discretized using the forward Euler method, to facilitate the design of the MPC controller. The discrete error model of the system is as follows:

$$\widetilde{\chi}(k+1) = \widetilde{a}\widetilde{\chi}(k) + \widetilde{b}\widetilde{u}(k) \tag{4}$$

where $\widetilde{a} = \begin{bmatrix} 1 & 0 & -Tv_r\sin\phi_r \\ 0 & 1 & Tv_r\cos\phi_r \\ 0 & 0 & 0 \end{bmatrix}$, $\widetilde{b} = \begin{bmatrix} T\cos\phi_r & 0 \\ T\sin\phi_r & 0 \\ \frac{T\tan\delta_r}{L} & \frac{Tv_r}{L\cos^2\delta_r} \end{bmatrix}$, and T is the sampling time.

2.1.2. Dynamic Model

The dynamic model of rear-wheel-steering agricultural machinery is schematically depicted in Figure 2, taking into account the two degrees of freedom in the lateral and yaw directions.

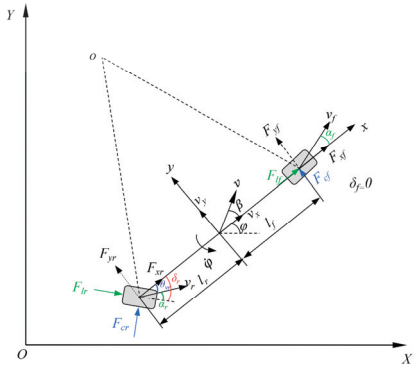


Figure 2. Schematic diagram of 2DoF dynamic model for rear-wheel-steering agricultural machinery. Note: F_{lf} and F_{lr} are the longitudinal forces acting on the front and rear wheels; F_{cf} and F_{cr} are the lateral forces acting on the front and rear wheels; F_{xf} and F_{xr} are the combined forces acting on the front and rear wheels along the x-axis; F_{yf} and F_{yr} are the combined forces acting on the front and rear wheels along the y-axis; v_x and v_y are the longitudinal and lateral velocities of the vehicle; l_f and l_r are the distances between the front and rear axles and the center of mass, respectively; θ_{vr} is the velocity deviation angle of the rear wheels.

The two-degree-of-freedom (2DoF) dynamic equation for rear-wheel-steering agricultural machinery is derived as follows:

$$m(\ddot{y} + v_x\dot{\phi}) = 2C_{cf}(-\frac{v_y + l_f\dot{\phi}}{v_x}) + 2C_{cr}(\delta_r - \frac{v_y - l_r\dot{\phi}}{v_x}) \tag{5}$$

$$I_z\ddot{\phi} = 2l_fC_{cf}(-\frac{v_y + l_f\dot{\phi}}{v_x}) - 2l_rC_{cr}(\delta_r - \frac{v_y - l_r\dot{\phi}}{v_x}) \tag{6}$$

where m is the mass of the vehicle; I_z is the moment of inertia around the z-axis of the vehicle; C_{cf} and C_{cr} are the lateral stiffness of the front and rear wheels respectively.

Choose $\chi = \left[y, \dot{y}, \phi, \dot{\phi}\right]^T$ as the state variables (i.e., lateral position, lateral position change rate, yaw angle, and yaw angle change rate). Set the rear-wheel angle as the control variable, i.e., $u = [\delta_r]$. The state space equation is as follows:

$$\frac{d}{dt}\begin{bmatrix} y \\ \dot{y} \\ \phi \\ \dot{\phi} \end{bmatrix} = \begin{bmatrix} 0 & 1 & 0 & 0 \\ 0 & -\frac{2C_{cf}+2C_{cr}}{mv_x} & 0 & -\frac{2C_{cf}l_f - 2C_{cr}l_r}{mv_x} - v_x \\ 0 & 0 & 0 & 1 \\ 0 & \frac{2l_rC_{cr} - 2l_fC_{cf}}{I_zv_x} & 0 & -\frac{2l_r^2C_{cr} + 2l_f^2C_{cf}}{I_zv_x} \end{bmatrix} \begin{bmatrix} y \\ \dot{y} \\ \phi \\ \dot{\phi} \end{bmatrix} + \begin{bmatrix} 0 \\ \frac{2C_{cr}}{m} \\ 0 \\ -\frac{2l_rC_{cr}}{I_z} \end{bmatrix} \delta_r \tag{7}$$

Rewrite Equation (7) as follows:

$$\dot{\chi} = a\chi + bu \tag{8}$$

The response of the vehicle's lateral displacement, yaw angle, lateral velocity, and yaw angular rate can be analyzed under a given steering angle input according to the system state equation. However, as the goal of tracking control is to minimize tracking deviation,

the state equation ought to be able to analyze the vehicle's tracking deviation under a given angle. Therefore, it is necessary to transform Equation (8) into the tracking error state space equation.

Set $\widetilde{\chi} = \left[\widetilde{y}, \dot{\widetilde{y}}, \widetilde{\phi}, \dot{\widetilde{\phi}}\right]^T$ as the state variables (i.e., lateral position error, lateral position error change rate, yaw angle error, and yaw angle error change rate). The tracking error state space equation of the dynamic model for rear-wheel-steering agricultural machinery is as follows:

$$\frac{d}{dt}\begin{bmatrix}\widetilde{y}\\ \dot{\widetilde{y}}\\ \widetilde{\phi}\\ \dot{\widetilde{\phi}}\end{bmatrix} = \begin{bmatrix}0 & 1 & 0 & 0\\ 0 & -\frac{2C_{cf}+2C_{cr}}{mv_x} & \frac{2C_{cf}+2C_{cr}}{m} & -\frac{2C_{cf}l_f-2C_{cr}l_r}{mv_x}\\ 0 & 0 & 0 & 1\\ 0 & \frac{2l_rC_{cr}-2l_fC_{cf}}{I_zv_x} & -\frac{2l_rC_{cr}-2l_fC_{cf}}{I_z} & -\frac{2l_f^2C_{cr}+2l_f^2C_{cf}}{I_zv_x}\end{bmatrix}\begin{bmatrix}\widetilde{y}\\ \dot{\widetilde{y}}\\ \widetilde{\phi}\\ \dot{\widetilde{\phi}}\end{bmatrix} + \begin{bmatrix}0\\ \frac{2C_{cr}}{m}\\ 0\\ -\frac{2l_rC_{cr}}{I_z}\end{bmatrix}\delta_r \quad (9)$$

Rewrite Equation (9) as follows:

$$\dot{\widetilde{\chi}} = \widetilde{a}\widetilde{\chi} + \widetilde{b}u \quad (10)$$

According to Equations (8) and (10), path tracking can be accomplished by utilizing MPC to obtain the appropriate amount of control at the lowest possible cost, hence lowering the tracking error.

The kinematic model is simple and practical, suitable for situations with low speeds and infrequent curvature changes. Nevertheless, it disregards the force interacting between wheels and ground, which can lead to side-slip at high speeds. The dynamic model considers the ground–wheel interactions and more vehicle physics, but it is more complex and requires balancing computational accuracy and real-time performance due to its difficult-to-access modeling parameters. This paper takes the kinematic model as the control model since agricultural machines have low speeds and complex ground conditions

2.2. Model Predictive Control

To obtain the optimal control sequence at the current moment, the MPC controller performs an optimization solution based on the prediction model, objective function, and constraint conditions. The principle can be summarized as predictive modeling, rolling optimization, and feedback correction.

2.2.1. Objective Function

To achieve a more stable path-tracking effect, the control quantity increments are constrained and new state variables are constructed.

$$\xi(k) = \left[\widetilde{\chi}(k) \quad \widetilde{u}(k-1)\right]^T \quad (11)$$

Substitute Equation (11) into the state space equation to obtain a new state space expression and output equation:

$$\begin{cases} \xi(k+1) = A\xi(k) + B\Delta\widetilde{u}(k) \\ \eta(k) = C\xi(k) \end{cases} \quad (12)$$

where $A = \begin{bmatrix}\widetilde{a} & \widetilde{b}\\ O_{N_u\times N_x} & I_{N_u}\end{bmatrix}$, $B = \begin{bmatrix}\widetilde{b}\\ I_{N_u}\end{bmatrix}$, $C = \begin{bmatrix}I_{N_x} & O_{N_u\times N_x}\end{bmatrix}$, $\Delta\widetilde{u}(k) = \widetilde{u}(k) - \widetilde{u}(k-1)$

N_x and N_u are the state and control quantity dimensions, respectively; I and O are the identity and zero matrix, respectively; and η is the new output variable.

Then the output equation can be expressed as follows:

$$Y = \Psi\xi(k) + \Theta\Delta U \quad (13)$$

where $Y = \begin{bmatrix} \eta(k+1) & \eta(k+2) & \cdots & \eta(k+N_p) \end{bmatrix}^T$, $\Psi = \begin{bmatrix} CA & CA^2 & \cdots & CA^{N_p} \end{bmatrix}^T$, $\Theta = \begin{bmatrix} CB & 0 & \cdots & 0 \\ CAB & CB & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ CA^{N_c-1}B & CA^{N_c-2}B & \cdots & CA^0B \\ \vdots & \vdots & \ddots & \vdots \\ CA^{N_p-1}B & CA^{N_p-2}B & \cdots & CA^{N_p-N_c}B \end{bmatrix}$, $\Delta U = \begin{bmatrix} \Delta\widetilde{u}(k) & \Delta\widetilde{u}(k+1) & \cdots & \Delta\widetilde{u}(k+N_c-1) \end{bmatrix}^T$

N_p and N_c ($N_p > N_c$) are the prediction and control horizon, respectively.

Thus, the output in N_p can be predicted if the present $\xi(k)$ and ΔU in the control horizon are known.

It is expected to achieve the reference value of the state variable as quickly as possible with the minimum control quantity while ensuring stationarity in path tracking control. Therefore, the optimization objective function can be expressed as follows:

$$\begin{aligned} J &= \widetilde{Y}^T Q_Q \widetilde{Y} + \Delta U^T R_R \Delta U \\ &= (Y - Y_r)^T Q_Q (Y - Y_r) + \Delta U^T R_R \Delta U \end{aligned} \tag{14}$$

where $Q_Q = I_{N_p} \otimes Q_{N_x \times N_p}$, $R_R = I_{N_p} \otimes R_{N_u \times N_c}$, Q and R are the state and quantity weight, respectively; and Y_r is the system output reference value, $Y_r = \begin{bmatrix} 0 & 0 & \cdots & 0 \end{bmatrix}^T$.

To facilitate the computational solution, the objective function is converted by the rounding method into the standard quadratic form:

$$\min_{\Delta U} J = \frac{1}{2}\Delta U^T H \Delta U + g^T \Delta U \tag{15}$$

where $H = \Theta^T Q_Q \Theta + R_R$, $g = E^T Q_Q \Theta$, $E = \Psi\xi(k)$. A relaxation factor ε is included to prevent the equation from having no solution, and the augmented vectors of H and g are as follows: $H = \begin{bmatrix} \Theta^T Q_Q \Theta + R_R & 0 \\ 0 & \varepsilon \end{bmatrix}$, $g = \begin{bmatrix} E^T Q_Q \Theta \\ \varepsilon \end{bmatrix}$.

2.2.2. Constraint Condition

The control quantities at each moment in the control horizon satisfy the following equations:

$$\begin{aligned} U = \begin{bmatrix} u(k) \\ u(k+1) \\ \vdots \\ u(k+N_c-1) \end{bmatrix} = \begin{bmatrix} u(k-1) \\ u(k-1) \\ \vdots \\ u(k-1) \end{bmatrix} + \begin{bmatrix} I_2 & 0 & \cdots & 0 \\ I_2 & I_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ I_2 & I_2 & \cdots & I_2 \end{bmatrix} \begin{bmatrix} \Delta u(k) \\ \Delta u(k+1) \\ \vdots \\ \Delta u(k+N_c-1) \end{bmatrix} = U_t + A_I \Delta U_t \\ U_{\min} = \begin{bmatrix} u_{\min} & u_{\min} & \cdots & u_{\min} \end{bmatrix}^T \le U \le \begin{bmatrix} u_{\max} & u_{\max} & \cdots & u_{\max} \end{bmatrix}^T = U_{\max} \\ \Delta U_{\min} = \begin{bmatrix} \Delta u_{\min} & \Delta u_{\min} & \cdots & \Delta u_{\min} \end{bmatrix}^T \le \Delta U \le \begin{bmatrix} \Delta u_{\max} & \Delta u_{\max} & \cdots & \Delta u_{\max} \end{bmatrix}^T = \Delta U_{\max} \end{aligned} \tag{16}$$

Rewrite Equation (16) as follows:

$$\begin{cases} A_I \Delta U_t \le U_{\max} - U_t \\ -A_I \Delta U_t \le -U_{\min} + U_t \\ \Delta U_{\min} \le \Delta U \le \Delta U_{\max} \end{cases} \tag{17}$$

For the standard quadratic programming problem, the MATLAB (2021b) QP solver "quadprog" is used for solving. Finally, a sequence of control increments in N_c is derived:

$$\Delta U_k^* = \begin{bmatrix} \Delta\widetilde{u}_k^* & \Delta\widetilde{u}_{k+1}^* & \cdots & \Delta\widetilde{u}_{k+N_c-1}^* \end{bmatrix}^T \tag{18}$$

Apply the first element of the control sequence to the controlled object, and the system continues to execute it until the next moment. Repeat the above operation in the next cycle.

$$u(k) = u(k-1) + \Delta\widetilde{u}_k^* \tag{19}$$

2.3. Analysis of Factors Influencing MPC Tracking Performance

It is vital to explore the impacts of the initial control parameters and the vehicle's original parameters on MPC since MPC is an optimization method, and the setting of its initial parameters has a significant impact on the solution results.

Typically, agricultural machinery operates at a speed of less than 3 m/s. The operating speed in the simulation below was set to 3 m/s, except for the speed impact analysis. Table 1 displays the other MPC control parameter settings, excluding the dynamically modified parameters. A MATLAB-based simulation environment was built for agricultural machinery path tracking control and the factors affecting tracking performance were analyzed.

Table 1. Control parameter settings.

Parameters	Values
Prediction horizon N_p	
Control horizon N_c	Dynamic adjustment
Preview value N_{pre}	
State quantity weight Q	100
Control increment weight R	1
Relaxation factor ε	10
Control minimum $\widetilde{u}_{\min}$/rad	$[-0.2 \quad -0.54]^T$
Control maximum $\widetilde{u}_{\max}$/rad	$[0.2 \quad 0.54]^T$
Control increment minimum $\Delta\widetilde{u}_{\min}$/rad	$[-0.05 \quad -0.2]^T$
Control increment maximum $\Delta\widetilde{u}_{max}$/rad	$[0.05 \quad 0.2]^T$
Sampling period T/s	0.1

2.3.1. Impact of Wheelbase on Tracking Effect

The wheelbase L is the only inherent physical property of the vehicle taken into account in the kinematic model. In agricultural machinery, the wheelbase is typically customized to a non-standard value according to the needs of specific applications. The following simulation analyzed the effect of a typical wheelbase on the effectiveness of path tracking, with settings N_p = 10 and N_c = 2.

As shown in Table 2 and Figure 3, with the same initial position and heading angle, the online distances are 8.7 m, 10.7 m, and 11.7 m, and the online times are 3 s, 3.6 s, and 3.9 s, respectively, when the wheelbases are 1.5 m, 2.9 m, and 3.7 m, respectively. That is to say, an increase in the wheelbase results in an increase in online distance and time, but the impact on the control accuracy and stability after going online is not significant. The heading error peak value before going online gradually decreases as the wheelbase increases, while the difference in rear-wheel steering angle peak value is not significant. This paper took the large wheelbase L = 3.7 m as the simulation parameter.

Table 2. Parameters of tracking effect at different wheelbases under straight paths.

Wheelbase /m	Online Time /s	Online Distance /m	Absolute Value of Lateral Error		
			Mean Value /m	Standard Deviation/m	Maximum Error/m
L = 1.5	3	8.7	0.0005	0.0027	0.0216
L = 2.9	3.6	10.7	0.0007	0.0029	0.0232
L = 3.7	3.9	11.7	0.0007	0.0031	0.0243

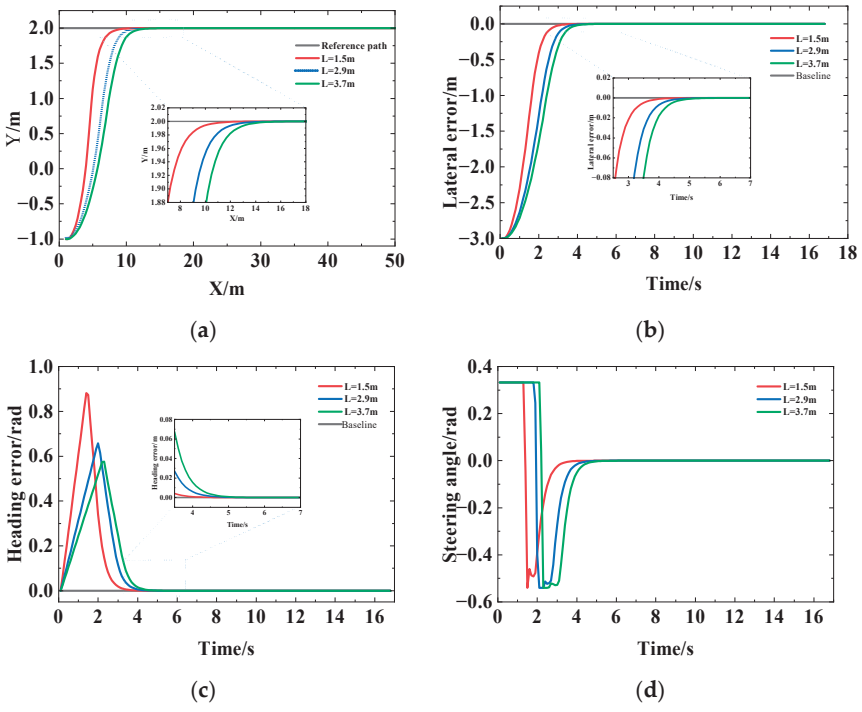


Figure 3. Tracking effects with different wheelbases under straight paths. (**a**) Path tracking; (**b**) lateral error; (**c**) heading error; (**d**) rear-wheel steering angle.

2.3.2. Impact of Preview on Tracking Effect

The gap from the actual position (x, y, φ) to the projected point of that position on the reference path is normally used as the value of the lateral error when path tracking computes it (Figure 4). Finding the nearest point to (x, y, φ) in the discrete reference path, i.e., the matching point $(x_m, y_m, \varphi_m, \rho_m)$, is the initial step in the calculation. It is usually assumed that the curvature of the matching point and the projection point is consistent, and then the projection substitution method is used to calculate approximately the lateral error e_d based on the matching point, i.e., as follows:

$$e_d \approx (\vec{x} - \vec{x}_m)\vec{n}_m \tag{20}$$

where $\vec{x}$ is the cartesian coordinate of the current position; $\vec{x}_m$ is the cartesian coordinates of the matching point; $\vec{x} - \vec{x}_m = (x - x_m, y - y_m)$, n_m, τ_m are the normal and tangential vectors of the matching point, respectively; and ρ_m is the curvature of the matching point.

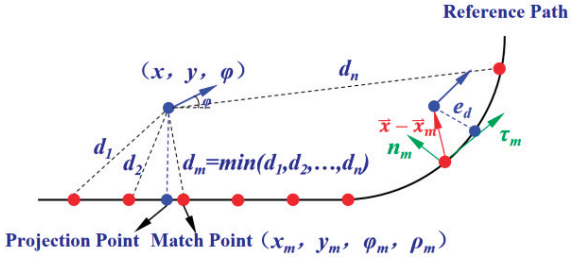


Figure 4. Schematic diagram of lateral error calculation.

This method is more accurate when the reference path has zero curvature. However, there is an inaccuracy in approximating the projection point with the matching point when the curvature is not zero. Consequently, adding a preview on the initial matching point is more beneficial to improving the tracking effect for path tracking with non-zero curvature, i.e., as follows:

$$Index_{Matchpoint} = Index[\min(d_1, d_2, \cdots d_n)] + N_{pre} \quad (21)$$

where *Index* represents the index number of the matching point; d_1, d_2, …, d_n represent the lengths from the vehicle and to each discrete point on the reference path, respectively; and N_{pre} stands for the added preview value, with a value range of [0, size (Reference Path) − N_{pre}].

1. Impact of reference path curvature on the value of N_{pre}

As shown in Figure 5, under the same other parameters (set N_p = 10, N_c = 3), adding N_{pre} is more effective when the reference path curvature is not zero. N_{pre} = 2 results in a much better curve tracking effect than N_{pre} = 0 and N_{pre} = 1, and when N_{pre} = 0, it is prone to untimely turn response with obvious hysteresis. However, when N_{pre} is excessively high (referring to N_{pre} = 5 in Figure 5), more attention will be paid to the distant situation, and ignoring the current situation will reduce control accuracy.

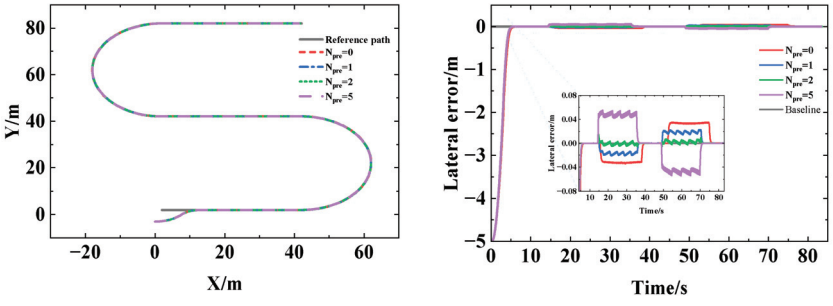


Figure 5. U-shaped path-tracking effect under different values of N_{pre}. (**a**) Path tracking; (**b**) lateral error.

From Figure 6 (set N_p = 10 and N_c = 3), N_{pre} should be increased appropriately but not excessively as the curvature grows. The tracking effect in Figure 6 is ranked as $N_{pre}(0) < N_{pre}(1) < N_{pre}(2) < N_{pre}(4) > N_{pre}(5) > N_{pre}(10)$, with N_{pre} = 4 having the best effect.

2. Impact of velocity on the value of N_{pre}

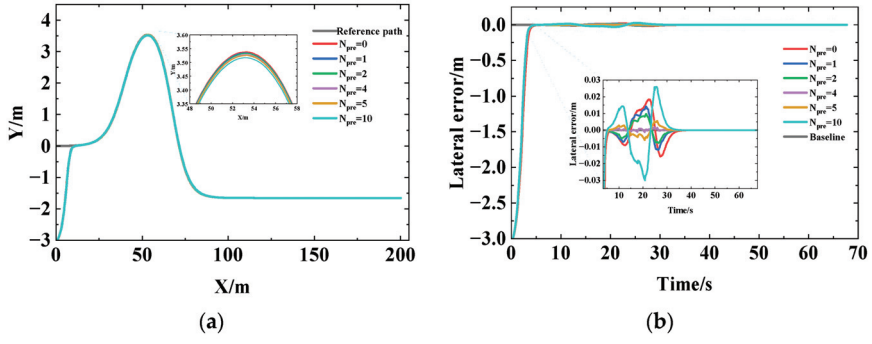


Figure 6. Effect of double-lane change path tracking at different values of N_{pre}. (**a**) Path tracking; (**b**) lateral error.

Take the U-shaped path (the commonly used agricultural machinery operation path) as the reference path and set N_p = 10 and N_c = 3.

Figure 7a illustrates that when v = 1 m/s, the tracking effect is ranked as $N_{pre}(0) < N_{pre}(1) > N_{pre}(2) > N_{pre}(3)$, where N_{pre} = 1 indicates the tracking effect's optimality. Figure 7b illustrates that when v = 3 m/s, the tracking effect is ranked as $N_{pre}(1) < N_{pre}(2) > N_{pre}(3) > N_{pre}(4)$, where N_{pre} = 2 indicates the tracking effect's optimality. Figure 7c illustrates that when v = 5 m/s, the tracking effect is ranked as $N_{pre}(2) < N_{pre}(3) > N_{pre}(4) > N_{pre}(5)$, where N_{pre} = 3 indicates the tracking effect's optimality. Figure 7d illustrates that when v = 10 m/s, the tracking effect is ranked as $N_{pre}(4) < N_{pre}(5) < N_{pre}(6) > N_{pre}(7)$, where N_{pre} = 6 indicates the tracking effect's optimality.

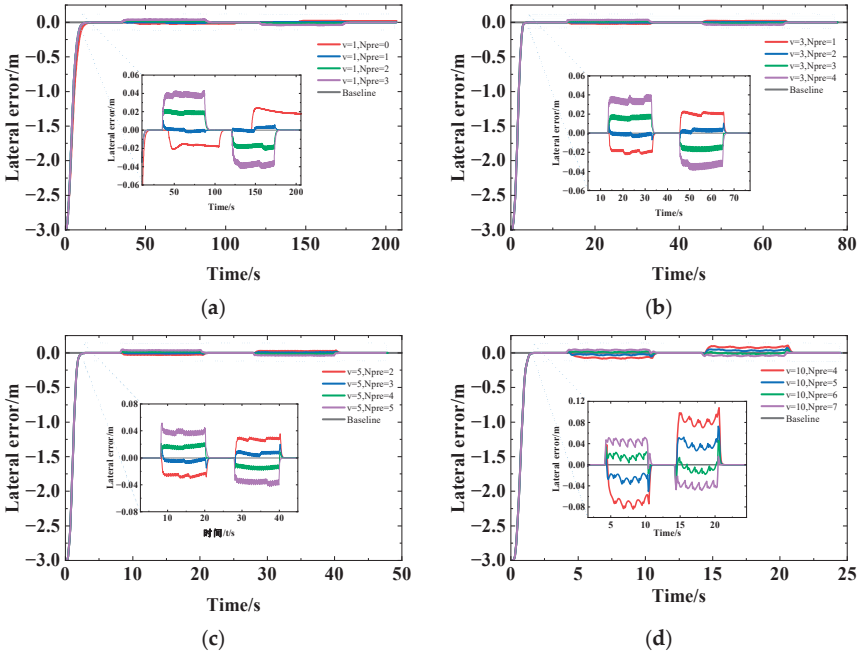


Figure 7. Lateral error under different N_{pre} values at different velocities. (**a**) Lateral error under different N_{pre} values at v = 1 m/s; (**b**) lateral error under different N_{pre} values at v = 3 m/s; (**c**) lateral error under different N_{pre} values at v = 5 m/s; (**d**) lateral error under different N_{pre} values at v = 10 m/s.

In summary, the selection of the value of N_{pre} is correlated positively with the curvature of the reference path and velocity.

2.3.3. Impact of the Predictive Horizon and Control Horizon on Tracking Effect

(1) When tracking a straight path (set N_{pre} = 0)

As can be seen in Figure 8a, a smaller N_p corresponds to slower convergence and a bigger N_p to faster online time. The difference in the control effect is not significant when N_c = 2, 3, and 4. Nevertheless, considering real-time control, N_p should be appropriately increased and N_c reduced while tracking a straight path; the tracking effect is best where N_p = 10 and N_c = 2, as shown in Figure 8. But once N_p rises to a certain point, convergence slows down while N_c remains constant, necessitating simultaneous adjustments to N_p and N_c, as Figure 8b illustrates. Thus, N_p should be adjusted adequately but not excessively for straight path tracking.

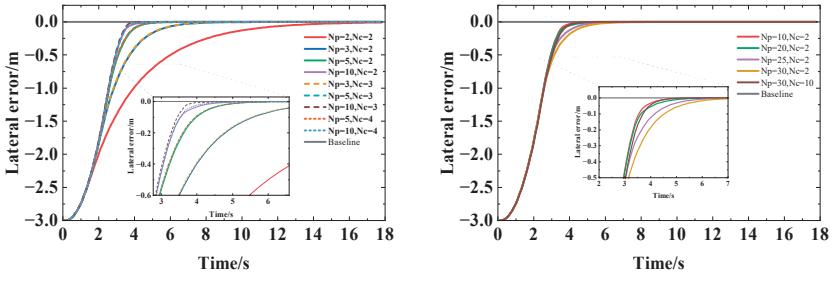


Figure 8. Lateral error under different N_p and N_c values of straight path tracking. (**a**) Lateral error I; (**b**) lateral error II.

(2) When tracking curves (set N_{pre} = 2)

In the case of the fixed-curvature circular path and the variable-curvature U-shaped path, as shown in Figures 9b and 10b, the control effect is better at N_c = 3 than it is at N_c = 2 when N_p = 3, 5, 10, respectively. While as N_c increases (referring to N_c = 5), the control impact does not significantly alter. As a result, N_c can be increased suitably for curve tracking while taking the computation of the real-time performance into account. When N_p = 10 and N_c = 3, there is a better effect, as shown in the illustration.

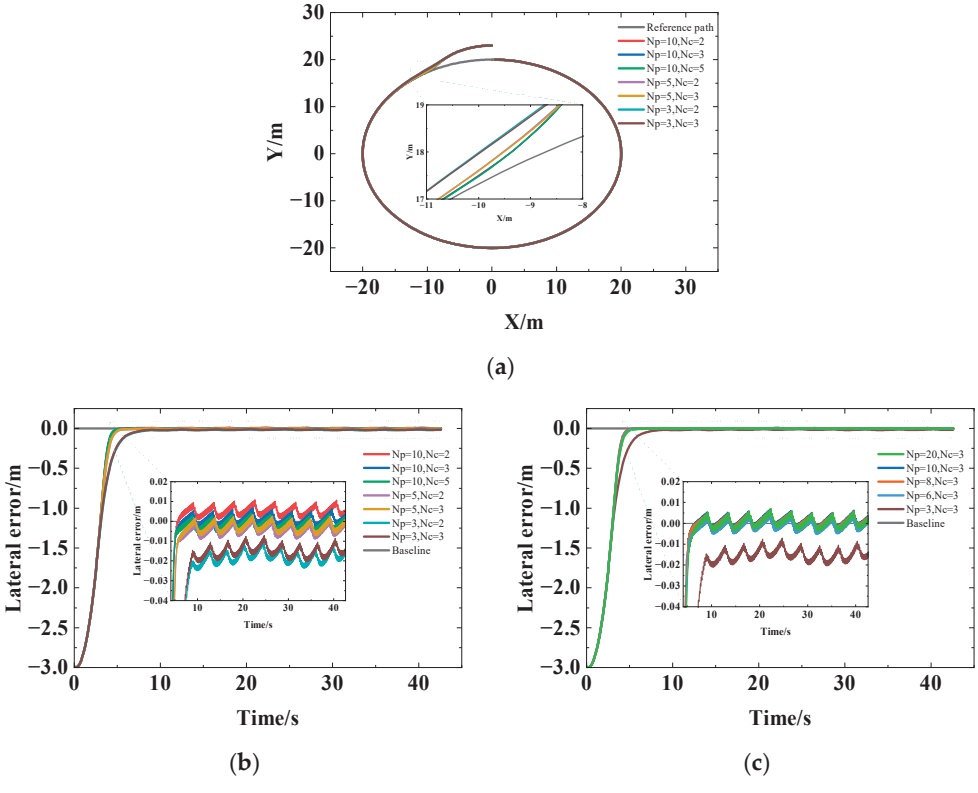


Figure 9. Circular path-tracking effect under different N_p and N_c values. (**a**) Path tracking; (**b**) lateral error I; (**c**) lateral error II.

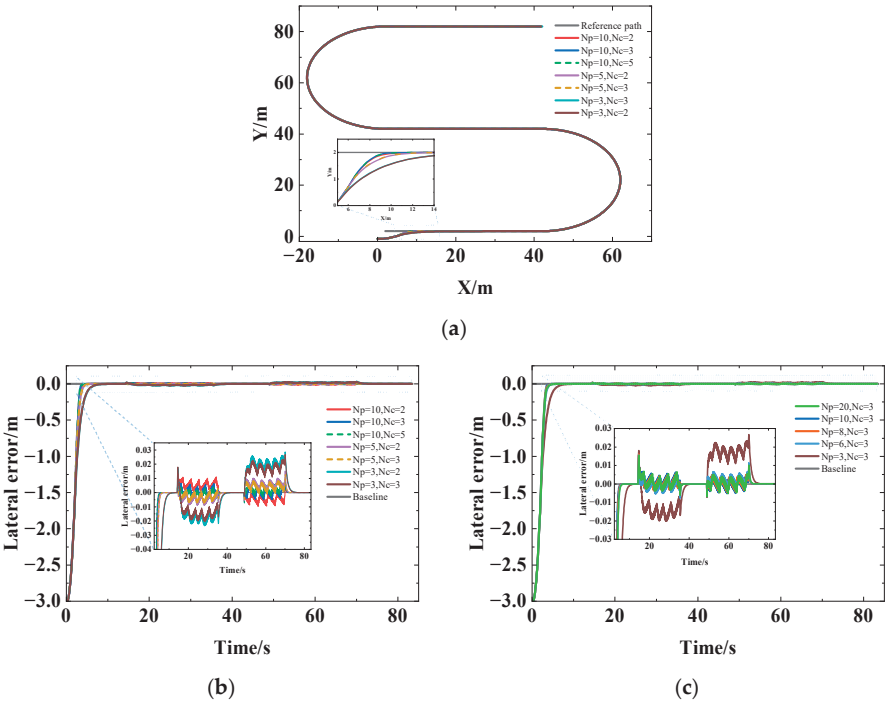


Figure 10. U-shaped path-tracking effect under different N_p and N_c values (**a**) Path tracking; (**b**) lateral error I; (**c**) lateral error II.

The difference in tracking effect is not significant when N_p = 6, 8, 10, 20, respectively, as shown by Figures 9c and 10c. Therefore, N_p can be suitably lowered considering the computational load and real-time computation when the curve is tracked. However, when it is reduced to a smaller value (referring to N_p = 3 and N_c = 3), the tracking effect is rather decreased.

In conclusion, the selection of N_p and N_c is affected by the reference path curvature ρ. N_p should be increased appropriately and N_c should be reduced appropriately when tracking a straight path. For curve tracking, N_p should be decreased appropriately to account for as much path curvature information as possible in each prediction horizon, and N_c should be increased appropriately to reduce the amplitude of changes in rear-wheel steering while ensuring tracking accuracy and stability.

At low vehicle speeds, the steering wheel needs to go a shorter distance to reach the target corner, resulting in a smaller total error during the process, while at high vehicle speeds, the vehicle will travel farther, potentially causing inferior tracking performance. Therefore, N_p should also grow when v increases to avoid the premature steering issue. N_c should also increase accordingly to reduce sudden changes in control inputs, preventing situations of skidding or even loss of control when driving at high speeds. The influence laws of reference path curvature ρ and vehicle speed v on N_p, N_c, and N_{pre} are summarized in Table 3.

It is obvious that varying operation speeds and path curvatures result in different tracking performances. The impact of changes in reference path curvature and agricultural machinery operating speed on tracking performance is not taken into account by the conventional MPC algorithm, which has fixed N_p, N_c, and N_{pre} values. N_p, N_c, and N_{pre} can be dynamically adjusted based on the reference path curvature and operating speed to achieve parameter adaptation.

Table 3. The influence laws of ρ and v on N_p, N_c, and N_{pre}.

Parameters		N_p	N_c	N_{pre}
ρ	↑	↓	↑	↑
	↓	↑	↓	↓
v	↑	↑	↑	↑
	↓	↓	↓	↓

Note: ↑ represents positive correlation; ↓ represents negative correlation.

2.4. Parameter Adaptive MPC

2.4.1. Parameter Adaptive MPC Based on Fuzzy Control

Fuzzy control (FC) has the advantage of handling nonlinear factors and the effects of uncertainty. The primary idea is to create fuzzy rules based on expert experience, then fuzzify the input, apply fuzzy reasoning, and defuzzify the results to obtain accurate outcomes. It benefits from being simple in design, easy to implement and transplant, and independent of the mathematical models of the controlled object [28]. Therefore, fuzzy control can be employed to perform the adaptive adjustment of N_p, N_c, and N_{pre}, based on the analysis of affecting factors in Section 2.3. Use v and ρ as the input parameters, while N_p, N_c, and N_{pre} are the outputs to the fuzzy control, i.e., as follows:

$$[N_p, N_c, N_{pre}] = Round[Fuzzy(v, \rho)] \tag{22}$$

where *Round* represents the rounding function, and *Fuzzy* represents a fuzzy control function.

Classify v, ρ, N_p, N_c, and N_{pre} into 7 fuzzy subsets, respectively: NB, NM, NS, ZO, PS, PM, and PB. An optimization strategy for output variables was developed based on the simulation results, and a fuzzy control rule library was established using the Mamdani method. Table 4 illustrates the fuzzy control rules and the range of values for each variable.

Table 4. Fuzzy control rules and range of variable values.

N_p = [5, 12]		v = [0, 3]						
		NB	NM	NS	ZO	PS	PM	PB
ρ = [0, 0.18]	NB	ZO	PS	PM	PM	PB	PB	PB
	NM	NS	ZO	PS	PM	PM	PB	PB
	NS	NM	NS	ZO	PS	PM	PM	PB
	ZO	NM	NM	NS	ZO	PS	PM	PM
	PS	NB	NM	NM	NS	ZO	PS	PM
	PM	NB	NB	NM	NM	NS	ZO	PS
	PB	NB	NB	NB	NM	NM	NS	ZO
N_c = [2, 5]		v = [0, 3]						
		NB	NM	NS	ZO	PS	PM	PB
ρ = [0, 0.18]	NB	NB	NB	NB	NM	NM	NS	ZO
	NM	NB	NB	NM	NM	NS	ZO	PS
	NS	NB	NM	NM	NS	ZO	PS	PM
	ZO	NM	NM	NS	ZO	PS	PM	PM
	PS	NM	NS	ZO	PS	PM	PM	PB
	PM	NS	ZO	PS	PM	PM	PB	PB
	PB	ZO	PS	PM	PM	PB	PB	PB
N_{pre} = [0, 4]		v = [0, 3]						
		NB	NM	NS	ZO	PS	PM	PB
ρ = [0, 0.18]	NB	NB	NB	NB	NM	NM	NS	ZO
	NM	NB	NB	NM	NM	NS	ZO	PS
	NS	NB	NM	NM	NS	ZO	PS	PM
	ZO	NM	NM	NS	ZO	PS	PM	PM
	PS	NM	NS	ZO	PS	PM	PM	PB
	PM	NS	ZO	PS	PM	PM	PB	PB
	PB	ZO	PS	PM	PM	PB	PB	PB

Figure 11 presents each variable's membership functions and the fuzzy control surface. Gaussian membership functions were applied at both ends of the input variable domain

and triangular membership functions were applied in the middle region of the domain to ensure high accuracy, stability, and sensitivity. The fuzzy outputs were obtained through fuzzy inference, and the precise output of N_p, N_c, and N_{pre} were obtained by using the MIN-MAX-Center of Gravity method for defuzzification.

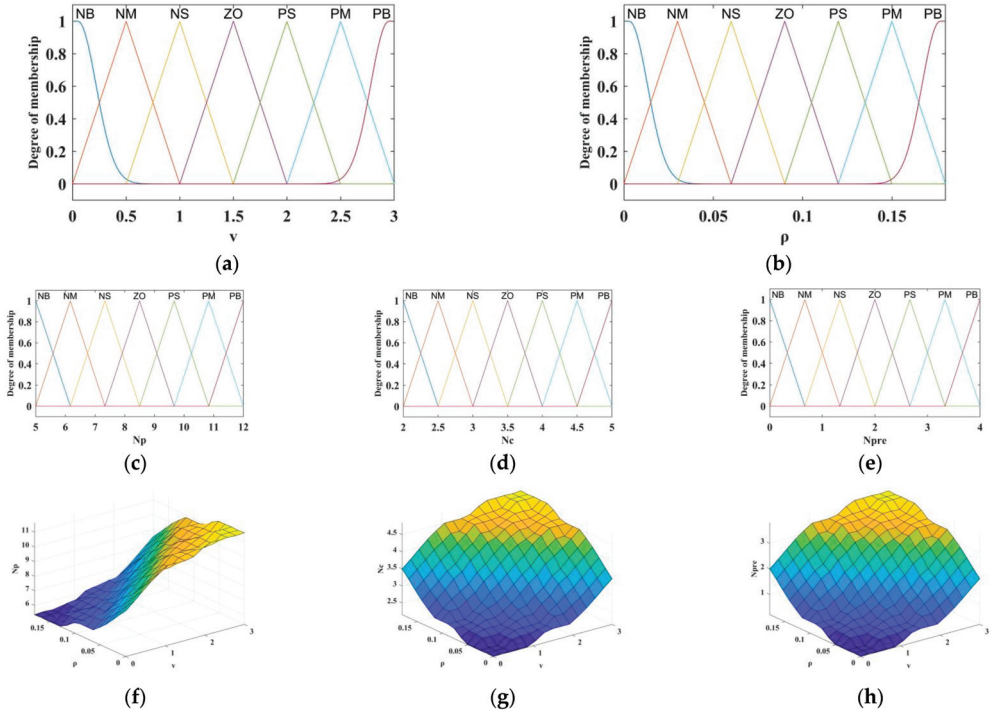


Figure 11. Relationship of input and output variables. (**a**) Membership function of v; (**b**) membership function of ρ; (**c**) membership function of N_p; (**d**) membership function of N_c; (**e**) membership function of N_{pre}; (**f**) fuzzy control surface of N_p; (**g**) fuzzy control surface of N_c; (**h**) fuzzy control surface of N_{pre}.

2.4.2. Parameter Adaptive MPC Based on PSO

Particle Swarm Optimization (PSO) is a heuristic global search algorithm based on swarm intelligence, characterized by its ease of implementation, high accuracy, and rapid convergence [29]. The basic idea involves leveraging collaboration and information exchange among particles to arrive at the optimal solution [30].

The PSO algorithm first initializes n particles to form a population $B = (B_1, B_2, \ldots, B_n)$, where every particle stands for a possible ideal solution. The position and velocity of the i-th particle Bi are $X_i = (x_{i1}, x_{i2}, \ldots, x_{id})$ and $V_i = (v_{i1}, v_{i2}, \ldots, v_{id})$, respectively. Then, the particles update their position and velocity by tracking individual extremum $P_{best} = (p_{i1}, p_{i2}, \ldots, p_{id})$ and population extremum $G_{best} = (p_{g1}, p_{g2}, \ldots, p_{gd})$ during the iteration process. The updated formula is as follows:

$$\begin{aligned} V_{id}^{k+1} &= \omega V_{id}^{k} + c_1 r_1 (P_{id}^{k} - X_{id}^{k}) + c_2 r_2 (P_{gd}^{k} - X_{gd}^{k}) \\ X_{id}^{k+1} &= X_{id}^{k} + V_{id}^{k+1} \end{aligned} \tag{23}$$

where ω is the inertia factor with a non-negative value; c_1 and c_2 are individual and group learning factors, respectively; and r_1 and r_2 are randomly generated numbers ranging from 0 to 1.

Local search benefits more from a smaller ω, while a bigger one is better for global search. Better optimization results can be achieved with the dynamic ω than with the fixed

one, as it can better balance the capabilities of local and global search. This paper adopted the commonly used Linear Decreasing Weight (LDW) strategy, i.e., as follows:

$$\omega_k = (\omega_{\text{int}} - \omega_{end})(G_k - k)/G_k + \omega_{end} \tag{24}$$

where G_k stands for the maximum number of iterations, k denotes the current iterations number, ω_{init} represents the initial inertia weight, and ω_{end} represents the inertia weight when evolving to the maximum iterations.

Set N_p, N_c, and N_{pre} as the three dimensions of each particle. Set the root mean square error of the lateral error as the fitness function, i.e., as follows:

$$Fitness(i) = MSE(e_d - r_{ed}) \tag{25}$$

where *Fitness* represents the fitness function, *MSE* represents the root mean square error function, and r_{ed} represents the lateral error reference value.

Table 5 displays the key parameter settings of the PSO algorithm, and Figure 12 illustrates the control process.

Table 5. PSO parameter settings.

Parameters	Values
Particle dimension D	3
Pop size	1
Maximum number of iterations	100
Individual learning factor c_1	0.8
Global learning factor c_2	1
VarSize of N_p	[5, 12]
VarSize of N_c	[2, 5]
VarSize of N_{pre}	[0, 4]
ω_{init}	0.9
ω_{end}	0.4

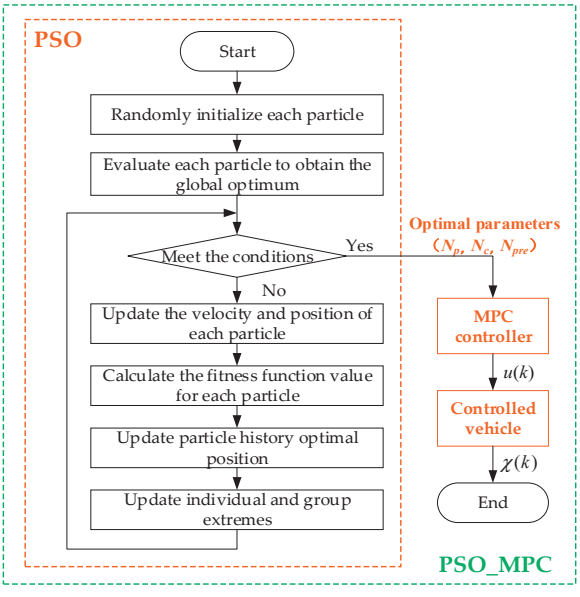


Figure 12. PSO-MPC control flow.

3. Results and Discussion

3.1. Simulation Test Results

The tracking effects of the manual tuning MPC (MPC), the parameter adaptive MPC based on fuzzy control (FC_MPC), and the parameter adaptive MPC based on PSO (PSO_MPC) under U-shaped paths (commonly used operation paths) and complex curve paths (simulating curve driving and obstacle avoidance behaviors, etc.) were compared and analyzed through simulation.

3.1.1. U-Shaped Path Simulation Test

Set $N_p = 6$, $N_c = 3$, and $N_{pre} = 2$ for the manual tuning MPC. From Figure 13 and Table 6, the tracking effects of MPC, FC_MPC, and PSO_MPC differ little and all have superior tracking effects under no external disturbance. In contrast, the effect of PSO_MPC is slightly better than the first two. The mean absolute values of lateral errors for the three methods are 0.0016 m, 0.0014 m, and 0.0003 m, respectively, with maximum values of 0.0238 m, 0.0014 m, and 0.0168 m. The mean absolute values of heading errors for the three methods are 0.0096 rad, 0.0095 rad, and 0.0095 rad, respectively, with maximum values of 0.0325 rad, 0.0321 rad, and 0.0326 rad.

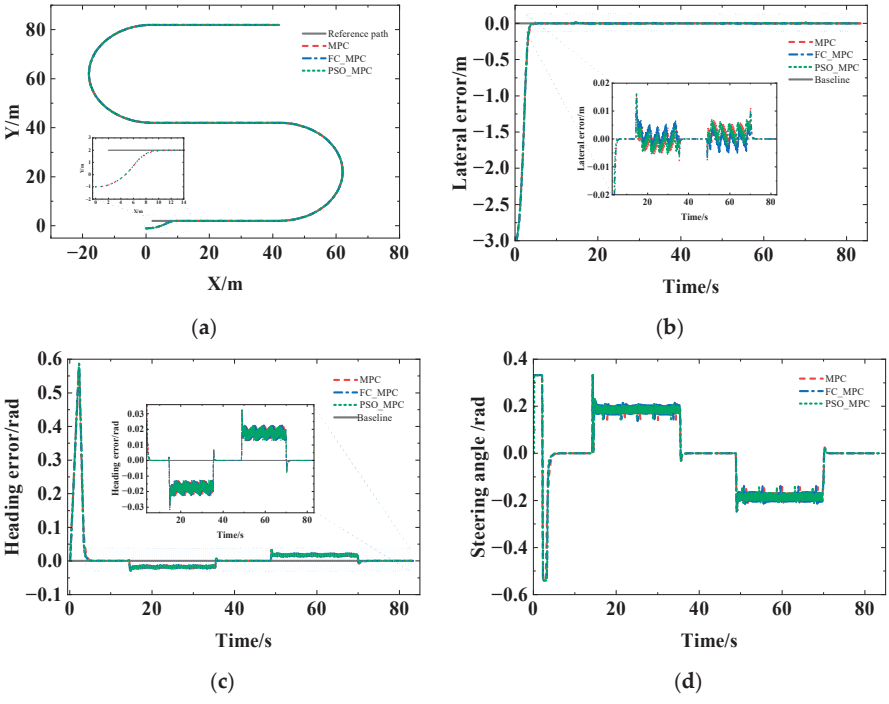
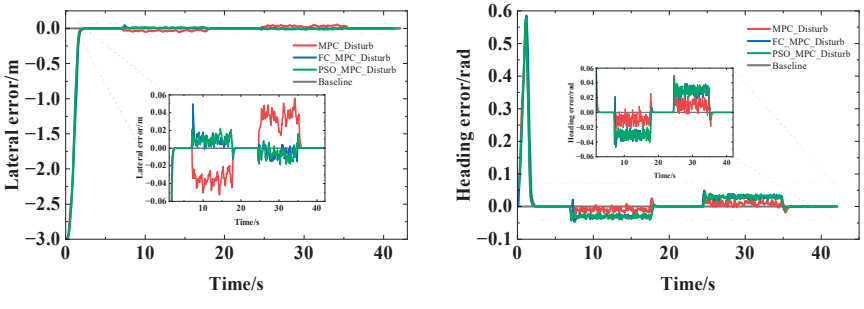


Figure 13. Contrast in tracking effects under a U-shaped path without disturbances. (**a**) Path tracking; (**b**) lateral error; (**c**) heading error; (**d**) rear-wheel steering angle.

Table 6. Statistics of lateral and heading errors under U-shaped path tracking without disturbances.

Control Algorithm	Online Time /s	Online Distance/m	Absolute Value of Lateral Error			Absolute Value of Heading Error		
			Mean Value /m	Standard Deviation /m	Maximum Error /m	Mean Value /rad	Standard Deviation /rad	Maximum Error /rad
MPC	4.1	11.29	0.0016	0.0023	0.0238	0.0096	0.0091	0.0325
FC_MPC	3.9	10.67	0.0014	0.0021	0.0214	0.0095	0.0091	0.0321
PSO_MPC	3.7	10.10	0.0003	0.0028	0.0168	0.0095	0.0091	0.0326

The tracking effects of FC_MPC and PSO_MPC are significantly better than that of MPC when the random perturbation of [1, 5] is set for the speed, as shown in Figure 14 and Table 7. There is a small difference between the two effects of FC_MPC and PSO_MPC, and the tracking effect is lower than when there is no perturbation but still within the usage demand range. The mean absolute values of lateral errors for the three methods are 0.0189 m, 0.0047 m, and 0.0049 m, respectively, with maximum values of 0.0561 m, 0.0498 m, and 0.0244 m. The mean absolute values of heading errors for the three methods are 0.0060 rad, 0.0160 rad, and 0.0159 rad, respectively, with maximum values of 0.0502 rad, 0.0471 rad, and 0.0463 rad.



(**a**) (**b**)

Figure 14. Comparison of tracking effects under a U-shaped path with disturbances. (**a**) Lateral error; (**b**) heading error.

Table 7. Statistics of lateral and heading errors under U-shaped path tracking with disturbances.

Control Algorithm	Online Time /s	Online Distance /m	Absolute Value of Lateral Error			Absolute Value of Heading Error		
			Mean Value /m	Standard Deviation /m	Maximum Error /m	Mean Value /rad	Standard Deviation /rad	Maximum Error /rad
MPC_Disturb	2.1	11.67	0.0189	0.0181	0.0561	0.0060	0.0076	0.0502
FC_MPC_Disturb	1.9	9.93	0.0047	0.0059	0.0498	0.0160	0.0153	0.0471
PSO_MPC_Disturb	1.7	10.44	0.0049	0.0056	0.0244	0.0159	0.0151	0.0463

3.1.2. Complex Curve Path Simulation Test

Set N_p = 10, N_c = 3, and N_{pre} = 2 for the manual tuning MPC. As can be observed from Figure 15 and Table 8, similar to the U-shaped path, MPC, FC_MPC, and PSO_MPC differ little and all have good tracking effects under no external disturbance. The mean absolute values of lateral errors for the three methods are 0.0017 m, 0.0018 m, and 0.0019 m, respectively, with maximum values of 0.0269 m, 0.0315 m, and 0.0290 m. The mean absolute values of heading errors for the three methods are 0.0021 rad, 0.0022 rad, and 0.0022 rad, respectively, with maximum values of 0.0334 rad, 0.0334 rad, and 0.0330 rad.

Similar to the U-shaped path, the tracking effects of FC_MPC and PSO_MPC are significantly better than that of MPC when the random perturbation of [−5, 5] is set for the speed, as shown by Figure 16 and Table 9. Furthermore, the difference in tracking performance between FC_MPC and PSO_MPC is relatively small. The mean absolute values of lateral errors for the three methods are 0.0072 m, 0.0020 m, and 0.0018 m, respectively, with maximum values of 0.1625 m, 0.0200 m, and 0.0193 m. The mean absolute values of heading errors for the three methods are 0.0022 rad, 0.0020 rad, and 0.0021 rad, respectively, with maximum values of 0.0293 rad, 0.0200 rad, and 0.0263 rad.

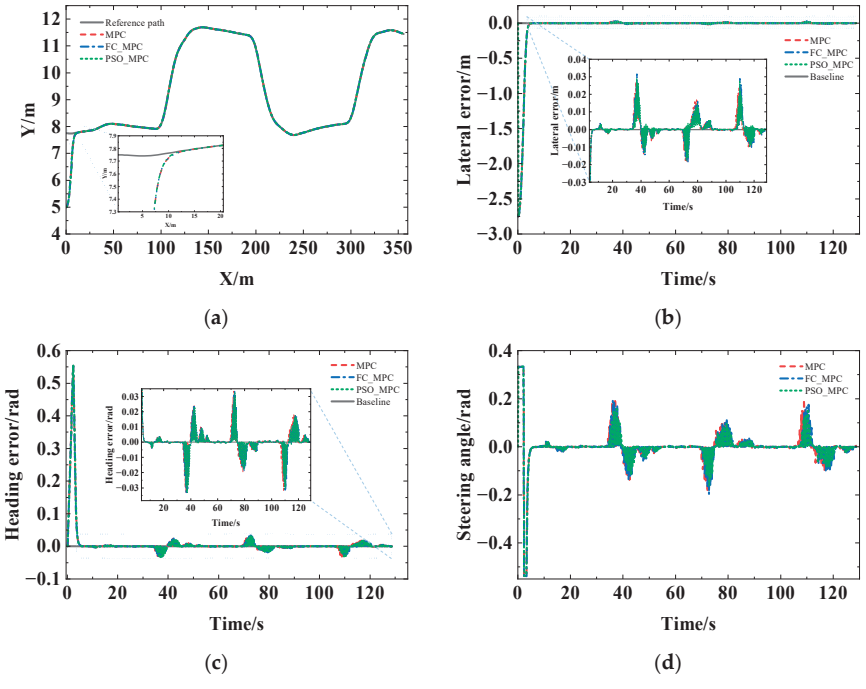


Figure 15. Contrast in tracking effects under complex curve path without disturbances. (**a**) Path tracking; (**b**) lateral error; (**c**) heading error; (**d**) rear-wheel steering angle.

Table 8. Statistics of lateral and heading errors under complex curve path tracking without disturbances.

Control Algorithm	Online Time /s	Online Distance/m	Absolute Value of Lateral Error			Absolute Value of Heading Error		
			Mean Value /m	Standard Deviation /m	Maximum Error /m	Mean Value /rad	Standard Deviation /rad	Maximum Error /rad
MPC	4.2	10.57	0.0017	0.0035	0.0269	0.0021	0.0046	0.0334
FC_MPC	4.3	10.81	0.0018	0.0037	0.0315	0.0022	0.0046	0.0334
PSO_MPC	4.2	10.45	0.0016	0.0034	0.0290	0.0022	0.0046	0.0330

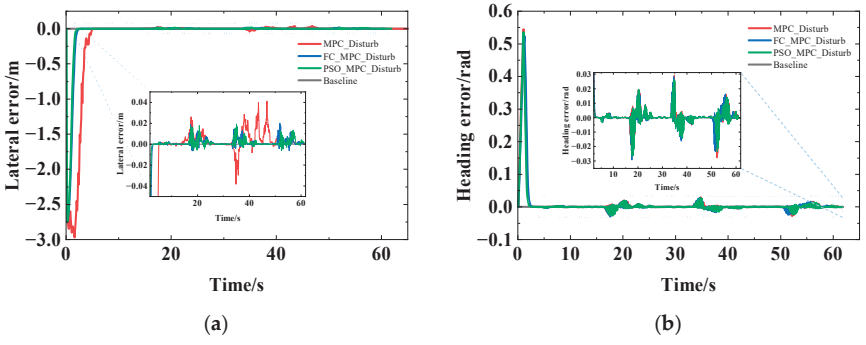


Figure 16. Contrast in tracking effects under complex curve path with disturbances. (**a**) Lateral error; (**b**) heading error.

Table 9. Statistics of lateral and heading errors in complex curve path tracking with disturbances.

Control Algorithm	Online Time /s	Online Distance/m	Absolute Value of Lateral Error			Absolute Value of Heading Error		
			Mean Value /m	Standard Deviation /m	Maximum Error /m	Mean Value /rad	Standard Deviation /rad	Maximum Error /rad
MPC_Disturb	4.9	26.59	0.0072	0.0137	0.1625	0.0022	0.0046	0.0293
FC_MPC_Disturb	2.3	11.12	0.0020	0.0033	0.0200	0.0020	0.0033	0.0200
PSO_MPC_Disturb	2.3	13.14	0.0018	0.0031	0.0193	0.0021	0.0044	0.0263

3.2. Discussion

According to the above simulation and analysis results, it can be seen that the three designed controllers, the manual tuning MPC, the FC_MPC, and the PSO_MPC, can achieve a better control effect under no disturbance, whether in the U-shaped path or complex curve path, meeting the requirements of ISO12188-2:2012(E) [31] and GB/T 37164-2018 [32] for agricultural machinery operation with a lateral error of less than 2.5 cm. This also demonstrates that, similar to PID parameter tuning, the MPC with manual parameter tuning under constant debugging can obtain the optimal key parameters combination, and that the control results satisfy the demand for control accuracy. Nevertheless, the debugging process is laborious, and the acquisition of debugging results has significant randomness and uncertainty. What is more, it lacks robustness and is not flexible enough to adjust to changing operating conditions. Due to the complex and ever-changing working environment of agricultural machinery, the FC_MPC and PSO_MPC with parameter adaptation designed in this paper have stronger applicability and application value.

The parameter adaptive MPC in this paper is based on the vehicle kinematic model, which is simple to control and can meet operational requirements, but it does not consider the effects of vehicle forces and ground conditions. Therefore, it is of certain research value and significance to design the parametric adaptive MPC controller based on the dynamic model of rear-wheel-steering vehicles (refer to Section 2.1.2), to compare and analyze the similarities and differences between the two in terms of the control effect, and to focus on summarizing the applicability conditions of the two, so as to make the controller's control accuracy, stability, and adaptability higher.

4. Conclusions

To enhance the precision and adaptability of path tracking, based on establishing kinematic and dynamic models of rear-wheel-steering agricultural machinery and analyzing the factors affecting the tracking effect, this study proposed the adaptive parameters MPC based on fuzzy control and PSO. The main conclusions are as follows:

- Since the vehicle model is a vital part of implementing MPC, the kinematic and dynamic error state–space equations for rear-wheel-steering agricultural machinery were established, which can directly apply to the design of MPC controllers.
- The factors impacting the MPC control effect were simulated and analyzed using the kinematic model as the control model. (a) The larger the unique vehicle intrinsic physical quantity wheelbase incorporated by the kinematic model, the larger the online time and distance. (b) Increasing N_{pre} was favorable to improving the curve tracking effect, and N_{pre} was correlated positively with the curvature and speed changes. Preferred values of N_{pre} for the U-shaped path were 1, 2, 3, and 6, respectively, at speeds of 1 m/s, 3 m/s, 5 m/s, and 10 m/s. (c) The tracking effect was affected by the parameter settings of N_p and N_c, which were influenced by the curvature and speed. N_p should be increased appropriately and N_c should be decreased appropriately in straight path tracking, while N_p should be decreased appropriately and N_c should be increased appropriately in curve tracking. N_p and N_c should be increased appropriately when the vehicle speed increases.
- Fuzzy control and the PSO algorithm were used to establish the adaptive MPC parameters (N_p, N_c, and N_{pre}) under different curvatures and velocities, and then the

simulation platform was built based on MATLAB for simulation and analysis under U-shaped and complex curve paths. The results indicated that the differences among the manual tuning MPC, the FC_MPC, and the PSO_MPC were small under no disturbance, and the tracking effects were all better. The mean absolute value of lateral error was $\leq$0.0018 m, with the maximum error <0.0315 m, while the mean absolute value of heading error was $\leq$0.0096 rad, with the maximum error <0.0325 rad. Laterally, this implies that manual tuning can obtain an optimal parameters combination, but with high uncertainty and low efficiency. The tracking effect of FC_MPC and PSO_MPC was significantly better than that of the manual tuning MPC under random velocity disturbance, and the difference between FC_MPC and PSO_MPC was not significant. As a result, FC_MPC and PSO_MPC are more anti-interference compared to the manual tuning MPC with fixed horizon, and more adaptable to complex field scenarios.

Author Contributions: Conceptualization, M.W. and X.T.; writing—original draft preparation, M.W. and X.T.; software, M.W.; validation, M.W. and C.N.; writing—review and editing, C.N., Z.W., Y.J. and J.J. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by the Key Research and Development Program Project of Xinjiang Uygur Autonomous Region (2022B02022-4), and the Central Guidance Local Science and Technology Development Special Fund Project (ZYYD2023C08).

Institutional Review Board Statement: Not applicable.

Data Availability Statement: The data presented in this study are available on request from the corresponding author.

Conflicts of Interest: The authors declare no conflicts of interest.

References

1. Luo, X.; Liao, J.; Hu, L.; Zhou, Z.; Zhang, Z.; Zang, Y.; Wang, P.; He, J. Research progress of intelligent agricultural machinery and practice of unmanned farm in China. *J. South China Agric. Univ.* **2021**, *42*, 8–17+5.
2. Meng, Z.; Wang, H.; Fu, W.; Liu, M.; Yin, Y.; Zhao, C. Research Status and Prospects of Agricultural Machinery Autonomous Driving. *Trans. Chin. Soc. Agric. Mach.* **2023**, *54*, 1–24.
3. Macenski, S.; Singh, S.; Martín, F.; Ginés, J. Regulated Pure Pursuit for Robot Path Tracking. *Auton. Robot.* **2023**, *47*, 685–694. [CrossRef]
4. Wang, L.; Zhai, Z.; Zhu, Z.; Mao, E. Path Tracking Control of an Autonomous Tractor Using Improved Stanley Controller Optimized with Multiple-Population Genetic Algorithm. *Actuators* **2022**, *11*, 22. [CrossRef]
5. Han, G.; Fu, W.; Wang, W.; Wu, Z. The Lateral Tracking Control for the Intelligent Vehicle Based on Adaptive PID Neural Network. *Sensors* **2017**, *17*, 1244. [CrossRef] [PubMed]
6. Lee, K.; Jeon, S.; Kim, H.; Kum, D. Optimal Path Tracking Control of Autonomous Vehicle: Adaptive Full-State Linear Quadratic Gaussian (LQG) Control. *IEEE Access* **2019**, *7*, 109120–109133. [CrossRef]
7. Li, J.; Shang, Z.; Li, R.; Cui, B. Adaptive Sliding Mode Path Tracking Control of Unmanned Rice Transplanter. *Agriculture* **2022**, *12*, 1225. [CrossRef]
8. Wang, M.; Chen, J.; Yang, H.; Wu, X.; Ye, L. Path Tracking Method Based on Model Predictive Control and Genetic Algorithm for Autonomous Vehicle. *Math. Probl. Eng.* **2022**, *2022*, e4661401. [CrossRef]
9. Arifin, B.; Suprapto, B.Y.; Prasetyowati, S.A.D.; Nawawi, Z. Steering Control in Electric Power Steering Autonomous Vehicle Using Type-2 Fuzzy Logic Control and PI Control. *World Electr. Veh. J.* **2022**, *13*, 53. [CrossRef]
10. Li, H.; Gao, F.; Zuo, G. Research on the Agricultural Machinery Path Tracking Method Based on Deep Reinforcement Learning. *Sci. Program.* **2022**, *2022*, e6385972. [CrossRef]
11. Zhang, S.; Wang, H.; Chen, P.; Zhang, X.; Li, Q. Overview of the application of neural networks in the motion control of unmanned vehicles. *Chin. J. Eng.* **2022**, *44*, 235–243.
12. A Survey on Industry Impact and Challenges Thereof [Technical Activities]. *IEEE Control Syst.* **2017**, *37*, 17–18. [CrossRef]
13. Liu, J.; Yang, Z.; Huang, Z.; Li, W.; Dang, S.; Li, H. Simulation Performance Evaluation of Pure Pursuit, Stanley, LQR, MPC Controller for Autonomous Vehicles. In Proceedings of the 2021 IEEE International Conference on Real-time Computing and Robotics (RCAR), Xining, China, 15–19 July 2021; pp. 1444–1449.
14. Diaz-del-Rio, F.; Sanchez-Cuevas, P.; Iñigo-Blasco, P.; Sevillano-Ramos, J.L. Improving Tracking of Trajectories through Tracking Rate Regulation: Application to UAVs. *Sensors* **2022**, *22*, 9795. [CrossRef]

15. Xue, P.; Wu, Y.; Yin, G.; Liu, S.; Shi, Y. Path Tracking of Orchard Tractor Based on Linear Time-Varying Model Predictive Control. In Proceedings of the 2019 Chinese Control And Decision Conference (CCDC), Nanchang, China, 3–5 June 2019; pp. 5489–5494.
16. He, J.; Hu, L.; Wang, P.; Liu, Y.; Man, Z.; Tu, T.; Yang, L.; Li, Y.; Yi, Y.; Li, W.; et al. Path Tracking Control Method and Performance Test Based on Agricultural Machinery Pose Correction. *Comput. Electron. Agric.* **2022**, *200*, 107185. [CrossRef]
17. Zhou, B.; Su, X.; Yu, H.; Guo, W.; Zhang, Q. Research on Path Tracking of Articulated Steering Tractor Based on Modified Model Predictive Control. *Agriculture* **2023**, *13*, 871. [CrossRef]
18. Wang, H.; Liu, B.; Ping, X.; An, Q. Path Tracking Control for Autonomous Vehicles Based on an Improved MPC. *IEEE Access* **2019**, *7*, 161064–161073. [CrossRef]
19. Shi, P.; Chang, H.; Wang, C.; Ma, Q.; Zhou, M. Research on Path Tracking Control of Autonomous Vehicles Based on PSO-BP Optimized MPC. *Automob. Technol.* **2023**, *7*, 38–46.
20. Guan, L.; Liao, P.; Wang, A.; Shi, L.; Zhang, C.; Wu, X. Path tracking control of intelligent vehicles via a speed-adaptive MPC for a curved lane with varying curvature. *Proc. Inst. Mech. Eng. Part D J. Automob. Eng.* **2024**, *238*, 802–824. [CrossRef]
21. Liang, J.; Zhu, F.; Cai, Y.; Chen, X.; Chen, L. Intelligent Vehicle Path Tracking Control Based on Complex Curvature Variation. *Automot. Eng.* **2021**, *43*, 1771–1779.
22. Lin, X.; Tang, Y.; Zhou, B. Improved Model Predictive Control Path Tracking Strategy Based on an Online Updating Algorithm With Cosine Similarity and a Horizon Factor. *IEEE Trans. Intell. Transp. Syst.* **2022**, *23*, 12429–12438. [CrossRef]
23. Dai, C.; Zong, C.; Chen, G. Path Tracking Control Based on Model Predictive Control With Adaptive Preview Characteristics and Speed-Assisted Constraint. *IEEE Access* **2020**, *8*, 184697–184709. [CrossRef]
24. Choi, Y.; Lee, W.; Yoo, J. A Variable Horizon Model Predictive Control Based on Curvature Properties of Vehicle Driving Path. *Trans. KSAE* **2021**, *29*, 1147–1159. [CrossRef]
25. Liu, H.; Sun, J.; Cheng, K.W.E. A Two-Layer Model Predictive Path-Tracking Control With Curvature Adaptive Method for High-Speed Autonomous Driving. *IEEE Access* **2023**, *11*, 89228–89239. [CrossRef]
26. Jing, L.; Zhang, Y.; Shen, Y.; He, S.; Liu, H.; Cui, Y. Adaptive Trajectory Tracking Control of 4WID High Clearance Unmanned Sprayer. *Trans. Chin. Soc. Agric. Mach.* **2021**, *52*, 408–416.
27. Shen, Y.; Zhang, Y.; Liu, H.; He, S.; Feng, R.; Wan, Y. Research Review of Agricultural Equipment Automatic Control Technology. *Trans. Chin. Soc. Agric. Mach.* **2023**, *54*, 1–18.
28. Belman-Flores, J.M.; Rodríguez-Valderrama, D.A.; Ledesma, S.; García-Pabón, J.J.; Hernández, D.; Pardo-Cely, D.M. A Review on Applications of Fuzzy Logic Control for Refrigeration Systems. *Appl. Sci.* **2022**, *12*, 1302. [CrossRef]
29. Gad, A.G. Particle Swarm Optimization Algorithm and Its Applications: A Systematic Review. *Arch. Comput. Methods Eng.* **2022**, *29*, 2531–2561. [CrossRef]
30. Li, D.; Wang, L.; Guo, W.; Zhang, M.; Hu, B.; Wu, Q. A Particle Swarm Optimizer with Dynamic Balance of Convergence and Diversity for Large-Scale Optimization. *Appl. Soft Comput.* **2023**, *132*, 109852. [CrossRef]
31. *ISO 12188-2:2012(E)*; Tractors and Machinery for Agriculture and Forestry-Test Procedures for Positioning and Guidance Systems in Agriculture-Part 2: Testing of Satellite-Based Auto-Guidance Systems during Straight and Level Travel. International Organization for Standardization: Geneva, Switzerland, 2012.
32. *GB/T 37164-2018*; Work Performance Requirements and Evaluation Method of Auto-Guidance Systems for Self-Propelled Agricultural Machinery. State Administration for Market Regulation. China National Standardization Administration: Beijing, China, 2018.

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.

agriculture

MDPI

Article

Improved Real-Time Models for Object Detection and Instance Segmentation for *Agaricus bisporus* Segmentation and Localization System Using RGB-D Panoramic Stitching Images

Chenbo Shi [1], Yuanzheng Mo [1], Xiangqun Ren [1], Jiahao Nie [1], Chun Zhang [1], Jin Yuan [2] and Changsheng Zhu [1,*]

[1] College of Intelligent Equipment, Shandong University of Science and Technology, Tai'an 271019, China; skd996523@sdust.edu.cn (C.S.); 202283230023@sdust.edu.cn (Y.M.); 202283230026@sdust.edu.cn (X.R.); 202283230040@sdust.edu.cn (J.N.); zhangchun@sdust.edu.cn (C.Z.)

[2] College of Mechanical and Electronic Engineering, Shandong Agricultural University, Tai'an 271018, China; jyuan@sdau.edu.cn

* Correspondence: zcs@sdust.edu.cn

Abstract: The segmentation and localization of *Agaricus bisporus* is a precondition for its automatic harvesting. *A. bisporus* growth clusters can present challenges for precise localization and segmentation because of adhesion and overlapping. A low-cost image stitching system is presented in this research, utilizing a quick stitching method with disparity correction to produce high-precision panoramic dual-modal fusion images. An enhanced technique called Real-Time Models for Object Detection and Instance Segmentation (RTMDet-Ins) is suggested. This approach utilizes SimAM Attention Module's (SimAM) global attention mechanism and the lightweight feature fusion module Space-to-depth Progressive Asymmetric Feature Pyramid Network (SPD-PAFPN) to improve the detection capabilities for hidden *A. bisporus*. It efficiently deals with challenges related to intricate segmentation and inaccurate localization in complex obstacles and adhesion scenarios. The technology has been verified by 96 data sets collected on a self-designed fully automatic harvesting robot platform. Statistical analysis shows that the worldwide stitching error is below 2 mm in the area of 1200 mm × 400 mm. The segmentation method demonstrates an overall precision of 98.64%. The planar mean positioning error is merely 0.31%. The method promoted in this research demonstrates improved segmentation and localization accuracy in a challenging harvesting setting, enabling efficient autonomous harvesting of *A. bisporus*.

Keywords: automatic picking robot; mushroom detection; attention mechanism; image processing; computer vision

Citation: Shi, C.; Mo, Y.; Ren, X.; Nie, J; Zhang, C.; Yuan, J.; Zhu, C. Improved Real-Time Models for Object Detection and Instance Segmentation for *Agaricus bisporus* Segmentation and Localization System Using RGB-D Panoramic Stitching Images. *Agriculture* **2024**, *14*, 735. https://doi.org/10.3390/agriculture14050735

Academic Editor: Gniewko Niedbała

Received: 28 March 2024
Revised: 5 May 2024
Accepted: 7 May 2024
Published: 9 May 2024

Copyright: © 2024 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (https://creativecommons.org/licenses/by/4.0/).

1. Introduction

A. bisporus, widely cultivated as an edible mushroom, is renowned for its delectable taste and rich nutritional content [1,2]. However, against continuous growth in large-scale production, harvesting *A. bisporus* remains labor-intensive and heavily dependent on skilled workers. Labor costs and harvesting efficiency emerge as primary constraints affecting productivity development [3]. Machine vision-based methods for the segmentation and localization of *A. bisporus* have gained widespread application in agricultural harvesting robot platforms, owing to their advantages such as automation, real-time capability, non-contact nature, and repeatability [4,5]. These methods are extensively helpful in agricultural harvesting robot platforms [6,7].

The accurate contour segmentation and precise identification of optimal harvesting points for *A. bisporus* are prerequisites for efficient harvesting. Segmentation and localization are typically handled by the visual module of the harvesting robot, equipped with an independent information system. However, in a factory-scale cultivation environment, the

narrow working space of single-layer mushroom beds makes it challenging to provide sufficient distance for obtaining a panoramic field of view. Particularly, the clustered growth habits of *A. bisporus* result in substantial occlusion, making it difficult for the visual system to accurately segment and locate the optimal harvesting points [8]. The ability to segment and locate in complex cultivation environments is a prerequisite for achieving efficient harvesting, and the development of machine vision brings hope to address this challenge.

Reed et al. [9] pioneered designing a mushroom harvesting robot based on two-dimensional vision in the United Kingdom. In laboratory conditions, they captured the positional information of mushrooms through cameras to guide the mechanical arm in completing the harvesting task. However, they did not address the overlapping issue in the *A. bisporus* case. Yu et al. [10] introduced a mushroom image region labeling technique based on sequential scanning, using numbers to label the central area of individual mushroom images, thereby achieving differentiation of mushroom images. However, this method merely allows rough segmentation, and lacks precision in localization. Refs. [11,12] proposed a preliminary segmentation algorithm and a novel iterative label generation method for initial watershed marking. Although these methods achieved a 95.7% accuracy in the effective recognition of *A. bisporus*, significant errors persisted in the radius fitting of the mushrooms. Furthermore, a method involving the computation of a global gradient threshold was introduced. It segmented the image based on the edge gradient features, generating a binary image that underwent filtering and morphological processing. This resulted in an effective recognition rate of over 96% for *A. bisporus*. However, the overall operational efficiency remained low due to the excessive computational workload [13].

Visual solutions based on RGB two-dimensional images encompass various mechanisms and technologies, but they struggle to attain high-precision depth information for the target. Unlike traditional RGB cameras, depth cameras employ an active imaging approach, providing depth information and proving more suitable for the complex lighting conditions in *A. bisporus* cultivation scenes than color cameras. Nathanael et al. [14] proposed an approach that initially segments using RGB information, then employs the Hough transform to determine the center and radius and, finally, utilized depth information to provide harvesting height for the mechanical arm. However, this method only utilizes depth information to determine the height coordinate, and does not enhance the accuracy of the segmentation and localization algorithm. According to recent research, a depth map stitching approach incorporating parallax correction and a layered watershed algorithm based on depth maps achieved a detection rate of 95.82% for occluded bilateral mushrooms in limited planting conditions. This approach may not correctly address complex planting conditions due to the simple dataset employed. However, the picture acquisition approach in the first phases has greatly inspired our research [15].

In recent years, deep learning technology has been widely used in the field of image segmentation. Compared to traditional instance segmentation methods, CondInst [16] simplifies the network architecture by dynamically generating a specific mask for each instance, eliminating the need for ROI cropping and feature alignment. It significantly simplifies the network structure and reduces the computational burden. SOLOv2 [17] introduces several innovations on the original SOLO instance segmentation network. These include using Matrix NMS to improve instances' positioning and segmentation accuracy. However, its performance could be improved when dealing with highly overlapping objects or extremely complex scenes. Zhong et al. [18] proposes a fully automated picking robotic system that is capable of planning and picking overlapping, dense, and discrete *A. bisporus* clusters in the field of view. The system was tested in a real *A. bisporus* factory, and the success rate of the robot picking *A. bisporus* was 94.1%. The above research shows that the deep learning model can handle complex image segmentation tasks in close to real-time, and has become the preferred technology for achieving efficient and accurate instance segmentation.

In summary, it has become increasingly challenging for traditional methodologies to achieve breakthroughs in the recognition of *A. bisporus*. The advantages of deep learning-

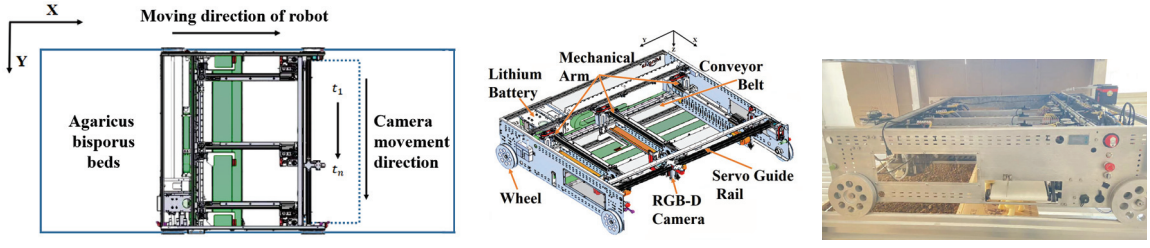
based segmentation algorithms are becoming increasingly prominent. Through analysis, it has been found that the majority of missed detections in current algorithms are due to occlusion of clusters of *A. bisporus* in complex environments. Building upon this premise, this paper proposes a high-precision segmentation and localization system for *A. bisporus* to enhance the algorithm's ability to extract features from small targets and address the challenge of accurately segmenting *A. bisporus* in complex environments. By incorporating a disparity-corrected fast stitching algorithm and an improved RTMDet-Ins instance segmentation network, the accuracy of segmentation and localization in complex cultivation environments is greatly enhanced. The key innovations and contributions of this research can be summarized as follows:

- To obtain panoramic information in low-visibility working environments, this paper introduces a cost-effective and high-precision image stitching method. The method utilizes multiple local color images and depth information with the fast-stitching algorithm to acquire a high-precision panoramic image through bimodal fusion.
- This paper proposes using SimAM's global attention mechanism and SPD-PAFPN's lightweight feature fusion module in the RTMDet-Ins instance segmentation network model to address the extensive occlusion caused by *A. bisporus*' clustered growth habits. This enhancement improves the handling capability of occluded *A. bisporus* and small-sized targets. Furthermore, the segmentation and detection capabilities of occluded areas for *A. bisporus* are further elevated by employing bimodal fusion images.

2. Materials and Methods

2.1. Harvest Robot Design

The harvesting robot for *A. bisporus* is equipped with three modules: visual, harvesting, and motion control. The visual module is the only way the harvesting robot perceives the external world, and gives the harvesting robotic arm positional information on the *A. bisporus*. Figure 1a shows the movement flow of the visual module. The harvesting module comprises servo motors, a robotic arm, flexible suction cups, and a conveyor belt. The motion control module coordinates duties for the visual and harvesting modules. Figure 1b,c depicts both the platform design and the physical machine.



(**a**) (**b**) (**c**)

Figure 1. The *A. bisporus* picking robot. (**a**) Movement flow of the visual module; (**b**) platform design; (**c**) physical machine.

The visual module undertakes the preliminary image processing tasks of segmentation and localization for harvesting, making the rational design of the visual module crucial for ensuring harvesting efficiency. The working distance for the camera to capture a panoramic image is insufficient due to the narrow operational space of mushroom beds. For instance, when the camera is mounted at a height corresponding to the current 300 mm layer, only a field of view measuring 400 mm × 300 mm is attainable, significantly smaller than the 15,000 mm × 1200 mm expanse of mushroom beds. Therefore, the image acquisition system is designed to operate in a parallel mode, capturing local sequential images and stitching them together to generate a composite panoramic image. Additionally, the dim planting environment and extensive occlusion from clustered mushroom groups make it challenging

for traditional color cameras to provide features and textures that meet the requirements of high-precision segmentation and localization.

The RealSense depth camera D435 (D435) made by American Intel Company (Santa Clara, CA, USA) is used to capture both RGB and depth images at the same time. The depth image is generated using laser speckle structured light imaging, where infrared laser emissions create diffraction patterns on the object's surface. The depth image provides harvesting height and compensates for the lack of color texture information in dim lighting conditions through depth information. Laser speckle typically offers higher texture resolution than Time-of-Flight imaging technology. The D435 camera has a significant cost advantage compared to line laser profile measurement devices. Moreover, the D435 camera provides a depth detection range of 0.2 m to 10 m, effectively covering the planting height of *A. bisporus*. Therefore, the low-cost D435 camera is chosen as the visual sensing device for the harvesting robot.

The camera is driven by a servo rail with a 0.02 mm offset error and an effective motion range of 1800 mm, ensuring its horizontal movement accuracy.

2.2. Visual Module Workflow Design

The visual module captures a sequence of images, and then stitches them into a panoramic image to complete an efficient image acquisition and processing workflow. Local images are captured during the slider's moving intervals. During the intervals when the slider is moving, local images are captured. Once the current scene capture is complete, the slider returns to the starting position. Simultaneously, the motor drives the robot to move along a horizontal rail along the edge of the mushroom bed, switching to the following local scene. The robot's movement and image-processing tasks occur concurrently.

The premise for segmentation and localization is the acquisition of high-precision global images. Building upon efficient data collection, we propose a high-precision fusion image stitching algorithm based on disparity correction. Initially, the collected sequences of color and depth images undergo denoising using the bilateral filtering algorithm. Subsequently, a bimodal image fusion operation is performed on the color and depth images. Following this, disparity correction is applied to the images, ultimately stitching the corrected fusion images into a panoramic image.

After obtaining the panoramic fusion image for the current region, the improved RTMDet-Ins model proposed in this paper is introduced to perform segmentation on the panoramic image. Subsequently, the segmentation results are fitted using the least squares method for ellipse fitting. Finally, the center of the fitted ellipse is taken as the picking point, guiding the robotic arm's operation with depth information. The algorithm processing flow is illustrated in Figure 2.

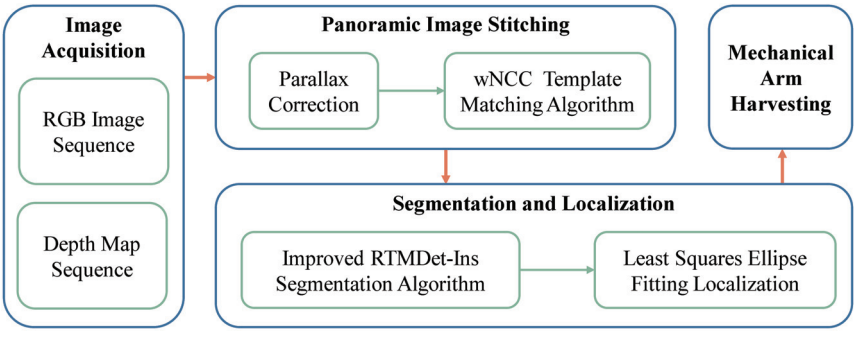


Figure 2. Algorithm processing flow.

2.3. High-Precision Fusion Image Stitching Algorithm Based on Vision Correction

Due to height restrictions between layers of mushroom beds, capturing a 1200 mm wide field of view in an instance is not feasible. Additionally, imaging at close distances

introduces significant disparities, severely impacting mushroom localization accuracy. Therefore, traditional stitching methods cannot meet the precision requirements. This paper proposes an RGB-D bimodal fusion image stitching algorithm based on disparity correction, aiming to achieve panoramic image stitching within the 1200 mm range.

2.3.1. Parallax Correction

In real-world planting scenes, there are *A. bisporus* with varying heights. Their projected positions on the image will differ because of different depths. Therefore, the transformation relationship between two *A. bisporus* at different depths in each image will inevitably differ. This disparity phenomenon can lead to ghosting artifacts in the stitched image, a crucial factor affecting stitching accuracy. This paper employs a correction method to mitigate stitching errors caused by disparities. As shown in Figure 3a, for two points A and B on *A. bisporus* with different depths and positions, when the camera receives the capture signal at time t_1, their projected coordinates on the imaging plane are A_1' and B_1', respectively. Similarly, at time t_2, the projected coordinates are A_2' and B_2'. The relative positional relationship between A and B at time t_1 is denoted as $P_{A_1'B_1'}$, and the relative positional relationship is denoted as $P_{A_2'B_2'}$.

By computing, we conclude that $P_{A_1'B_1'} \neq P_{A_2'B_2'}$, indicating the need to correct disparities before image stitching to eliminate the errors caused by this phenomenon. In fact, model analysis shows that the relative positional relationship of the orthographic projections of points A and B in the image is fixed and invariant, as shown in Figure 3b.

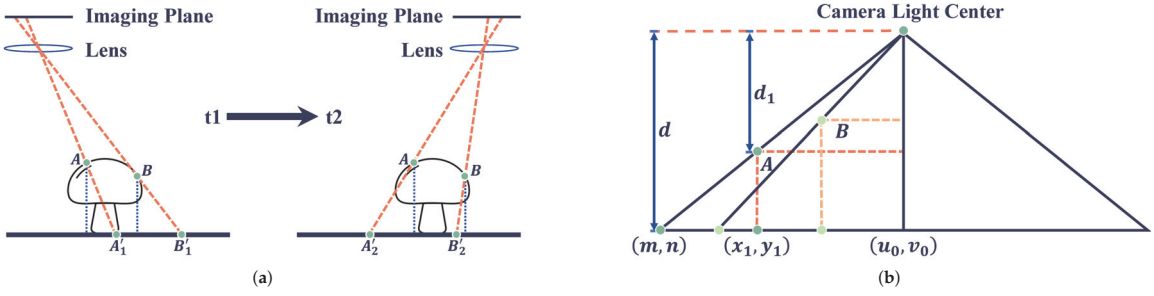


Figure 3. Parallax principle. (**a**) Generation process; (**b**) correction process.

The coordinates of point A, vertically projected onto the imaging plane in the image, are (x_1, y_1), and its distance relative to the camera plane is d_1. The coordinates of the camera's optical center on the image are (u_0, v_0), and the maximum depth value is represented by d. According to the principles of similar triangles, the following formulas can be derived:

$$\frac{u_0 - m}{d} = \frac{x_1 - m}{d - d_1} \tag{1}$$

$$\frac{v_0 - n}{d} = \frac{y_1 - n}{d - d_1} \tag{2}$$

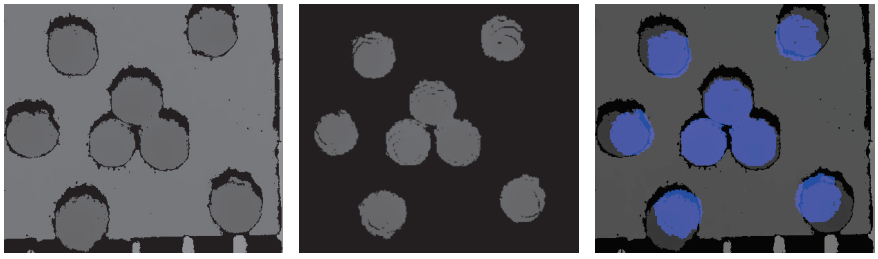
After rearranging, Equations (1) and (2) become:

$$x_1 = u_0 - \frac{d_1}{d}(u_0 - m) \tag{3}$$

$$y_1 = v_0 - \frac{d_1}{d}(v_0 - m) \tag{4}$$

A corrected local depth map is obtained by applying Equations (3) and (4) to the depth map of *A. bisporus*, and the existing disparities tend to disappear. The results after disparity correction are shown in Figure 4, where Figure 4a is without disparity correction, and the positions of *A. bisporus* are deviated from their actual locations. Figure 4b represents the

disparity correction algorithm after background separation, and the positions of *A. bisporus* are corrected. Figure 4c shows the before-and-after positional relationship by overlaying the original image with the corrected result.



(**a**) (**b**) (**c**)

Figure 4. Parallax correction results. (**a**) Original image; (**b**) parallax correction results after background separation; (**c**) positional relationship of *A. bisporus* before and after parallax correction.

After disparity correction, the locally corrected image will be synthesized into an RGB-D dual-modal fusion image. Dual-modal images are more suitable for high-precision segmentation, as they combine visual and spatial information. The color image conveys the texture and color details of the scene, while the depth image provides distance information for each pixel. This fusion method allows for a more accurate definition of object boundaries in the scene, thereby improving the accuracy of scene segmentation.

The D435 camera can simultaneously capture color and depth images, outputing images with registered states. The fusion of color and depth images is achieved by grayscale processing of both images and feeding them into their respective channels, as illustrated in Figure 5. Figure 5a represents the depth energy map, where the orange portion signifies compost, and the blue corresponds to *A. bisporus*. The varying saturation levels of blue denote the differences in elevation of the *A. bisporus*, with higher saturation indicating greater height.



(**a**) (**b**) (**c**)

Figure 5. Local deep images and color images of allocation and fusion results. (**a**) Deep energy map; (**b**) Color image; (**c**) Fusion image after the standard.

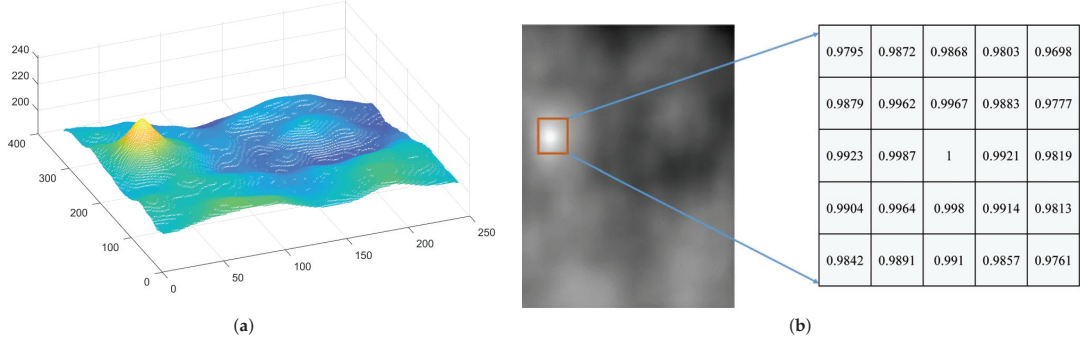
2.3.2. Panorama Stitching

Image matching constitutes a pivotal pre-splicing step to seek images or features most akin to the given image. In this context, it is crucial to determine the optimal overlapping position between two adjacent sequential images. The template matching algorithm utilizing normalized cross-correlation exhibits remarkable robustness, manifesting substantial insensitivity to noise and outliers. It concurrently ensures precision while enhancing the real-time nature of the matching process [19]. Consequently, this manuscript adopts the weighted Normalized Cross-Correlation ($wNCC$) template matching algorithm as the fundamental strategy for image stitching. The $wNCC$ template matching algorithm employs a sliding window to search the image to be registered, computing the similarity between the template image and the sub-image of identical dimensions. Considering the distinctive

features of the *A. bisporus* image, the formula introduces additional weighting coefficients, as expressed in Equation (5).

$$wNCC(x,y) = \frac{\sum_{i=1}^{w}\sum_{j=1}^{h} \xi(x,y)\left(I_{x,y}(i,j) - \overline{I}_{x,y}\right)\left(T_{x',y'}(i,j) - \overline{T_{x',y'}}\right)}{\sqrt{\sum_{i=1}^{w}\sum_{j=1}^{h}\left(I_{x,y}(i,j) - \overline{I_{x,y}}\right)^2}\sqrt{\sum_{i=1}^{w}\sum_{j=1}^{h}\left(T_{x',y'}(i,j) - \overline{T_{x',y'}}\right)^2}} \quad (5)$$

where I represents the matched image, T signifies the template image, (x,y) denotes the current matching position, and $\xi(x,y)$ represents the weighting coefficient. When (x,y) lies within the *A. bisporus* region, its grayscale value significantly surpasses that of the composting region. At this juncture, $\xi(x,y)$ is set to 1, whereas it is set to 0.1 when in a non-*A. bisporus* region. $\overline{I}$ denotes the grayscale average of the current search region, $\overline{T}$ represents the grayscale average of the template image, and $wNCC(x,y)$ signifies the degree of matching at the (x,y) position. The absolute value of the matching result never exceeds 1. When it equals to 1, it indicates the highest level of correspondence between the sub-image in the image to be registered and the template image, as illustrated in Figure 6a. Here, the X and Y axes denote the current coordinates, while the Z axis represents the computed result of $wNCC(x,y)$, with the yellow area indicating a high matching level. Figure 6b is the grayscale image, where the X and Y axes represent the current coordinates, and the brightness values signify the computed results of $wNCC(x,y)$. Brighter areas in the image denote higher matching degrees.



(**a**) (**b**)

Figure 6. *wNCC* Template matching algorithm principle. (**a**) 3D curved image; (**b**) gray image.

Given the corresponding states of color imagery and depth information, the stitching algorithm, aimed at reducing stitching time, computes the overlapping regions based on depth information and utilizes the computational results to merge the images seamlessly. In reality, achieving a flawless match during the matching process between adjacent images is not guaranteed on every occasion. In instances where multiple high-response regions emerge, employing color imagery for secondary verification facilitates the identification of the optimal matching window. Figure 7 illustrates a series of locally stitched images, applying a disparity correction algorithm to generate a panoramic depth map.

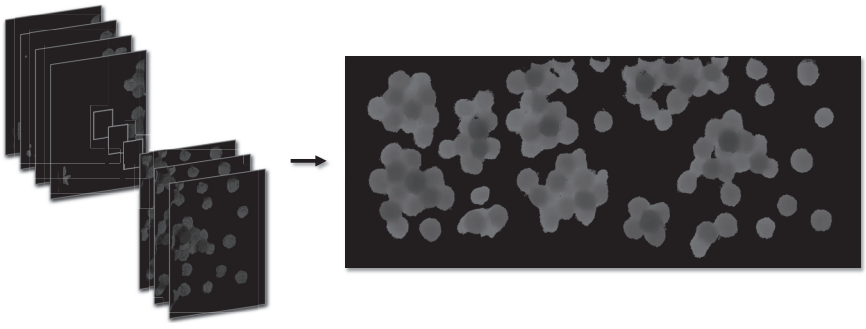


Figure 7. Panorama depth map synthesis results.

2.4. Improved RTMDet-Ins Fusion Image Experimental Segmentation and Localization Algorithm

RTMDet-Ins [20] is a single-stage target instance segmentation algorithm that comes with both performance and efficiency. It fully leverages the advantages of large-kernel depth-wise convolutions and a dynamic soft label assignment strategy, achieving a remarkable 44.6% Mask AP on the COCO dataset. However, the precision of *A. bisporus* segmentation has yet to meet the requirements for production scenarios. To further enhance segmentation precision, an analysis of segmentation failures was conducted. The clustered nature of *A. bisporus* led to mutual occlusion, resulting in incomplete mushroom caps in the images. Additionally, the misidentification of small-sized targets was induced by the chaotic planting environment. These factors emerge as the current algorithm's primary constraints in segmenting and detecting *A. bisporus*.

This study proposes an improved instance segmentation algorithm for *A. bisporus* based on RTMDet-Ins, addressing the abovementioned issues. It integrates the global attention mechanism from SimAM [21] and the SPD-PAFPN feature fusion module into the RTMDet-Ins instance segmentation algorithm. The improved overall network framework is illustrated in Figure 8.

The backbone network module primarily consists of the Convolutional module and Cross-Stage Partial SimAM Attention module (CSP-Sim). Precisely, the Convolutional module consists of a conv2d convolution, a BN layer, and a SiLU activation function. The CSP-Sim module is composed of three Convolutional modules, a CSP-Block with residual connections, and a SimAM attention module. Each CSP-Block in CSP-Sim is composed of one Convolutional module and one depthwise separable convolution [22] with a large kernel. This design structure helps to enhance the model's ability to extract and represent image features. After passing through the backbone network, the image generates three different scale feature maps, which are then sent to the feature fusion module for processing. The feature fusion module adopts the SPD-PAFPN module, which combines the Feature Pyramid Network (FPN) [23], the Path Aggregation Network (PAN) [24], and SPD-Conv [25]. This module innovatively reverses the conventional direction of the FPN feature pyramid and employs SPD-Conv for downsampling. It achieves effective fusion of multi-scale features by preserving feature information for small targets, while also integrating and reconstructing high-level and shallow-level features from diverse scale feature maps. The fused features are fed into the segmentation module, consisting of a kernel prediction head and a mask feature head, similar to CondInst. The mask feature head consists of four convolutional layers that extract mask features with eight channels from features at multiple levels. Meanwhile, the kernel prediction head forecasts a 169-dimensional vector for each instance. This vector is then decomposed into three dynamic convolution kernels. These kernels interact with the mask features and coordinate features to produce instance segmentation masks. We employ dice loss [26] as the supervision method for the instance masks, following standard practices.

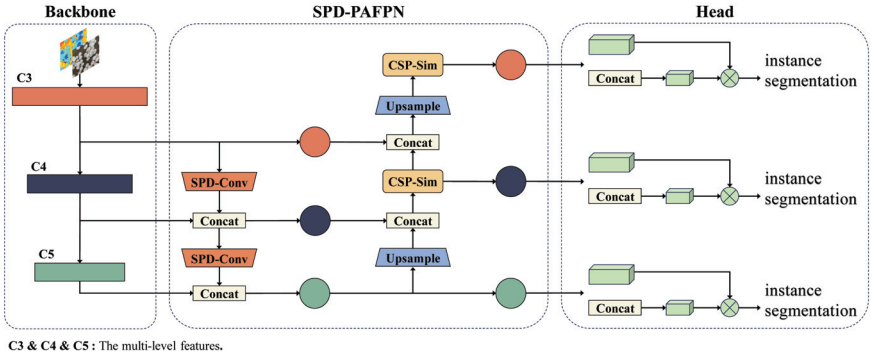


Figure 8. Improved RTMDet-Ins network architecture.

2.4.1. Improved to the Base Unit CSP-Sim

CSP-Sim [27] leveraged a cross-stage feature fusion strategy to enhance the variability of learned features in different layers, significantly reducing computational load and improving inference speed and accuracy. Attention mechanisms are employed to reinforce a deep learning model's extraction of crucial features while reducing the dispersion of target information. SimAM represents a mechanism where channel and spatial attention coexist, deriving genuine 3-D attention weights for feature maps without additional parameters. The mechanism enhances the model's capability to extract compelling features from RGB-D fused images. We have replaced the original attention mechanism in the fundamental building unit of CSP-Sim, and the modified CSPLayer is depicted in Figure 9.

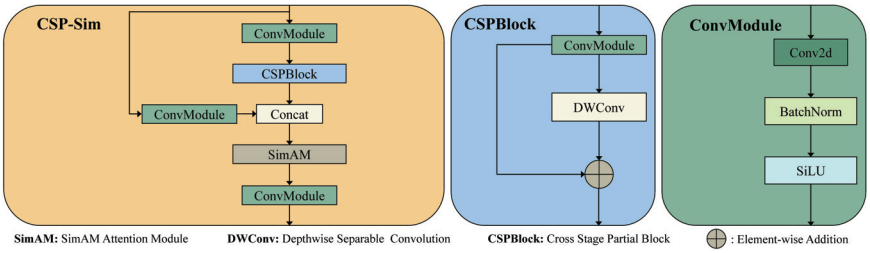


Figure 9. CSP-Sim Module Structure.

The SimAM attention mechanism, as compared to the original network, significantly enhances the ability to focus effectively on crucial regions while maintaining the same parameter scale. SimAM defines an energy function to measure the linear separability between an individual feature and all other features within the same channel to rapidly and efficiently locate important feature information. The energy function for each feature is defined as follows:

$$e_t(w_t, b_t, y, x_i) = \frac{1}{M-1} \sum_{i=1}^{M-1} (-1 - (w_t x_i + b_i))^2 + (1 - (w_i t + b_t))^2 + \lambda w_t^2 \tag{6}$$

where t and x_i represent the target feature and other features within the same channel, respectively. w_t and b_t denote the linear transformation weights and bias for t. The index i pertains to the spatial dimension, λ is a hyperparameter, and M is the number of all feature information in a single channel. The transformation weights and bias are represented as follows:

$$w_t = \frac{2(t - u_t)}{(t - u_t)^2 + 2\sigma_t^2 + 2\lambda} \tag{7}$$

$$b_t = -\frac{1}{2}(t + u_t)w_t \tag{8}$$

where $u_t = \frac{1}{M-1}\sum_{i=1}^{M-1} x_i$ and $\sigma_t^2 = \frac{1}{M-1}\sum_{i}^{M-1}(x_i - u_t)^2$ represent the mean and variance, respectively, of all feature information except t. By computing w_t and b_t along with the mean and variance of all feature information in the channel, the minimum energy formula is obtained as follows:

$$e_t^* = \frac{4(\hat{\sigma}^2 + \lambda)}{(t - \hat{\mu})^2 + 2\hat{\sigma}^2 + 2\lambda} \tag{9}$$

where $\hat{\mu} = \frac{1}{M}\sum_{i=1}^{M} x_i$ and $\hat{\sigma}^2 = \frac{1}{M}\Sigma_{i=1}^{M}(x_i - \hat{u}_t)^2$. Equation (9) indicates that the lower the energy function value e_t, the greater the distinction between the feature information and its surrounding features, making it more crucial for image processing. Therefore, the importance of each feature information can be obtained through $1/e_t$.

2.4.2. Improved Feature Fusion Module

Due to the downsampling process in the feature fusion of RTMDet-Ins, which reduces the size of feature maps, there is a loss of fine-grained information. The relatively lower feature extraction capability also rapidly declines detection accuracy at lower resolutions. To address this issue, SPD-Conv is employed in the feature fusion module to replace the original downsampling module. This substitution helps reduce the loss of detailed information while enriching feature details, preserving more effective information from the RGB-D fused image. SPD-Conv is a traditional image transformation technique applied within the CNN [28], composed of a space-to-depth layer and a non-stride convolutional layer. The process of SPD-Conv is illustrated in Figure 10. The leftmost and rightmost sections of the image represent the input image and the output image, respectively. The middle section displays different colors to denote different image channels.

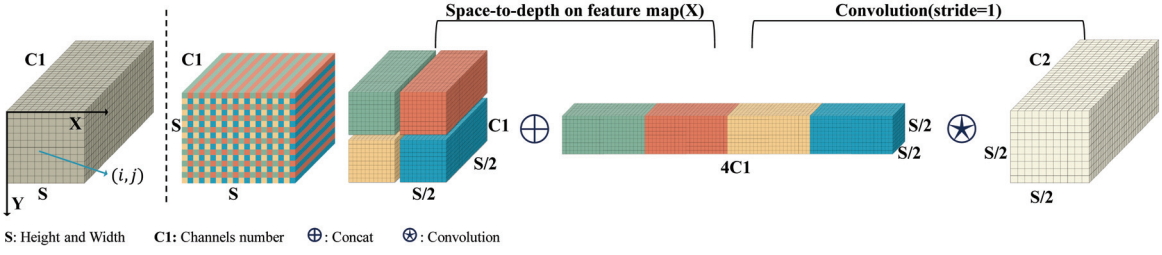


Figure 10. SPD-CONV process.

The steps are as follows: Firstly, the feature map $X(S \times S \times C_1)$ is split N times along the channel axis. Each resulting sub-feature map has dimensions of $\frac{S}{N}$ in length and width while maintaining the same channel count. These sub-feature maps are then merged along the channel axis to form the feature map X_1, where $X(S \times S \times C_1) \rightarrow X_1(\frac{S}{N} \times \frac{S}{N} \times N^2C_1)$. Subsequently, a non-stride convolutional layer ($Stride = 1$) is applied to the feature map X_1, resulting in the feature map X_2, where $X_1(\frac{S}{N} \times \frac{S}{N} \times N^2C_1) \rightarrow X_2(\frac{S}{N} \times \frac{S}{N} \times C_2)$.

Unlike other natural images, the distinctive aspect of the *A. bisporus* fused image is that in the presence of a more extensive background image, the targets to be detected are more diminutive and arranged more densely. As convolutional layers increase in the backbone network, the receptive field gradually enlarges, providing richer high-level instance information and leading to the loss of feature information for small-sized targets. However, for *A. bisporus* targets, the feature information for small-sized targets carried by shallow feature maps is more critical. The aim is to minimize unnecessary loss during the processing and increase the proportion of feature information for small-sized targets when fusing feature information.

The structure of the SPD-PAFPN module is depicted in the middle section of Figure 8. By reversing the traditional direction of the FPN feature pyramid, changing it from top-down to bottom-up, i.e., $(C5 > C4 > C3) \rightarrow (C3 > C4 > C5)$, early fusion from low-level occurs due to this alteration in the direction of the feature pyramid. Since downsampling is employed during this fusion to reduce the size, there is a risk of premature loss of feature information for small-sized targets in the fused image. SPD-Conv is utilized to downsample and address this. After the feature pyramid fusion, the feature information carried by the shallow feature map C3 remains unchanged. However, the deep feature maps $C4/C5$ have already incorporated the information from the shallow layers. When the subsequent path aggregation network adjusts its direction correspondingly from the bottom-up, it enables the flow of these previously fused shallow feature information from $C4/C5$ back to the shallow layers. The reflow increases the proportion of feature information for small-sized targets in the shallow feature map and incorporates high-level instance information.

2.4.3. Localization Algorithm Based on Least Squares Ellipse Fitting

Determining the segmentation center is a crucial step for precise localization. The mature contour of *A. bisporus* caps typically appears circular. However, in the case of partially inclined growth, the frontal projection of the contour may approximate an ellipse. Therefore, employing ellipse fitting is an ideal strategy for obtaining the segmentation center. Least squares fitting [29] balances efficiency and robustness, and can be integrated with segmentation results. Hence, this study chooses the least squares ellipse fitting to obtain the segmentation center.

2.4.4. Performance Indicators

The experiment adopts the Average Precision (AP) metric from the COCO dataset as the criterion to evaluate the algorithm's segmentation performance. The AP metric considers metrics such as Intersection over Union (IoU), *Precision* (P), and *Recall* (R) for both predicted and actual masks, providing a comprehensive reflection of the algorithm's overall performance. The P-value, Precision, is the ratio of true positives in the model's predicted data. The R-value, or Recall, is the ratio of true positives in the actual data. The specific mathematical expressions are as follows:

$$Precision = \frac{TP}{TP + FP} \tag{10}$$

$$Recall = \frac{TP}{TP + FN} \tag{11}$$

$$AP = \int_0^1 P(R)dR \tag{12}$$

where True Positive (TP) represents the number of correctly detected *A. bisporus*, False Positive (FP) denotes the number of falsely detected *A. bisporus*, and False Negative (FN) represents the number of missed *A. bisporus*. In addition to accuracy metrics for model detection, parameters (Params) and Floating Point Operations (FLOPs) are commonly used to evaluate model performance, reflecting its complexity. To provide a more intuitive representation of model performance, we also conducted statistics on the algorithm's runtime.

3. Results and Discussion

To comprehensively evaluate the various performance metrics of the proposed method in this paper, this section presents experiments on global stitching errors for the image stitching algorithm, ablation experiments for the image segmentation algorithm, cross-sectional comparative experiments for the segmentation algorithm, and comprehensive segmentation and localization performance experiments in complex environments.

3.1. Experimental Environment and Training Strategy

The RTMDet-Ins model's training environment was based on Python 3.7.16, PyTorch 1.7.0, and CUDA 11.0.221. It was operated on a server configured with an Intel (R) Core (TM) i7-12700k 3.6 GHz CPU, 32.00 GB RAM, and an NVIDIA (R) GeForce RTX (TM) 3090 (24 GB) GPU. The server ran on the Ubuntu 20.04 operating system (OS).

The dataset was collected from numerous clustered *A. bisporus* scenarios designed in the laboratory, totaling 1320 local images. Through stitching, 15 panoramic depth maps, panoramic color images, and panoramic fused images with dimensions of 1200 mm × 400 mm were obtained. The dataset comprises 1188 complete *A. bisporus* samples.

The training strategy utilizes the AdamW optimizer for iterative updates of network parameters, with a momentum parameter set to 0.05. The initial learning rate is set to $2 \times 0.004/(32 \times 8)$, the batch size is set to 2, and the number of iterations is set to 60. A linear warm-up learning rate is applied for the first 1000 batches, with an initial warm-up learning rate of 10^{-5}, gradually increasing to 2.5^{-4}. From the 30th iteration onward, a cosine annealing mechanism is employed, continuously decreasing the learning rate to 1.25^{-5} until the end of training.

3.2. Analysis of the Stitch Experiment Results

The primary sources of error in image stitching arise from misalignment in the stitching position and horizontal disparity caused by camera translation. The experiment quantifies the extent of error by comparing the distance difference between the stitched position of reference points and their actual positions. A total of 10 different scenarios, each featuring the cultivation of various *A. bisporus*, were employed for the stitching tests. This study required a more significant number of local images than conventional approaches to ensure image accuracy, with 44 local images needed for each scene and 43 stitching iterations. The global error is measured as the Euclidean distance between the measured coordinates of markers placed at the scene's tail and their actual coordinates, as presented in the test data shown in Table 1.

Table 1. Panoramic stitching error results.

Sample ID	No. 1	No. 2	No. 3	No. 4	No. 5	No. 6	No. 7	No. 8	No. 9	No. 10
Panoramic Stitching Errors after Disparity Correction/mm	1.855	1.932	2.021	1.903	2.029	1.892	1.923	1.929	2.004	1.969
Panoramic Stitching Errors without Correction/mm	15.236	17.768	18.234	16.246	18.675	19.023	15.912	16.824	18.897	17.932

The experimental results indicate that the average error decreases to 1.9457 mm after undergoing rectification through orthographic projection. Compared to the unprocessed image stitching results, the corrected error accounts for only 11.13% of the uncorrected error. Transforming the images into a three-dimensional orthographic view for stitching and fusion significantly reduces errors compared to unprocessed stitched images, aligning with practical application precision.

3.3. Ablation Study

To validate the segmentation and detection efficacy of the proposed algorithm in complex scenarios, we conducted ablation experiments on various improvement modules using the original RTMDet-Ins network as a baseline while maintaining the environment and parameters as constants. The symbol "✓" indicates the addition of the corresponding module to the model, and bold font signifies the optimal results for each column, as illustrated in Table 2. The experimental results of RTMDet-Ins are presented in the first row of Table 2, serving as the benchmark for other improvement experiments.

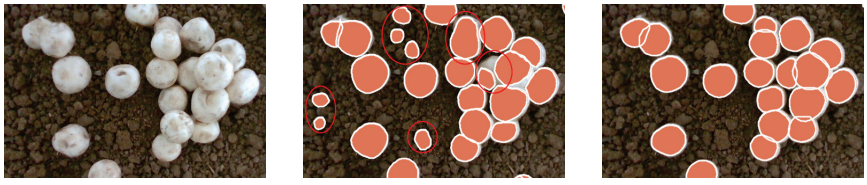
RTMDet + SimAM involves optimizing the fundamental construction unit CSP-Sim and introducing the SimAM global attention mechanism to replace the original attention mechanism, increasing 2.0% and 2.62% in AP^{50} and AP^{75}, respectively. RTMDet + SPD-CONV

improves the feature fusion module SPD-PAFPN, improving accuracy while reducing the parameter count by 1.4%. Integrating both improvement points in RTMDet-Ins leads to a 4.63% reduction in parameters, and the running time has been reduced to 25.38 ms while maintaining detection accuracy at 98.50% and 96.10%. Ablation experiments confirm each improvement point's effectiveness and compatibility, validating the enhancement approach's rationality.

Table 2. Improved point ablation experiment.

Methods	SimAM	SPD-Conv	Params (M)	FLOPs (G)	Running Time (ms)	AP^{50} (%)	AP^{75} (%)
RTMDet-Ins			5.61	11.87	26.26	96.50	93.48
RTMDet-Ins + SimAM	✓		5.46	11.87	25.73	97.91	95.73
RTMDet-Ins + SPD-Conv		✓	5.53	**10.74**	26.15	96.93	94.20
Improved RTMDet-Ins	✓	✓	**5.35**	**10.74**	**25.38**	**98.50**	**96.10**

The segmentation effects are shown in Figure 11. Figure 11b presents the segmentation results of RTMDet-Ins, where multiple cases of mis-segmentation of small targets and incorrect segmentation of occluded *A. bisporus* can be observed in the red circled parts. In contrast, as shown in Figure 11c, our method benefitted from the newly added SimAM attention module and SPD-Conv module, which correctly segmented all targets. Therefore, from a qualitative and quantitative perspective, the effectiveness of our method can be demonstrated.



(**a**) (**b**) (**c**)

Figure 11. Segmentation results. (**a**) Original image; (**b**) RTMDet-Ins; (**c**) our method.

3.4. Analysis of Detection Experiment Results

We randomly selected ten complex scenarios from the dataset to validate the enhanced performance of the RTMDet-Ins algorithm in the authentic segmentation of dense clusters of *A. bisporus*. On average, each scenario consists of dense clusters of *A. bisporus*, constituting approximately 70% of the scene.

To further validate the superiority of our algorithm, we compared it with other segmentation algorithms, including SOLOv2 [17], CondInst [16], classic deep learning-based instance segmentation algorithms, and the traditional watershed algorithm [15]. The experimental results are shown in Table 3.

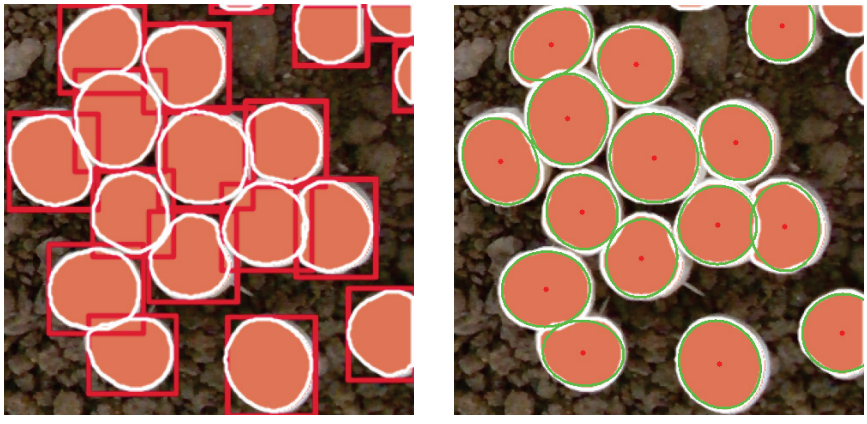
The data in Table 3 indicates that our method achieved an average accuracy of 98.64%, reaching 100% in various scenarios. Compared with deep learning algorithms such as CondInst and SOLOv2, the average accuracy increased by 1.78% and 2.27%, respectively. The average accuracy increased by 6.42% compared to the traditional watershed algorithm. Our analysis of error segmentation cases in other networks and traditional methods found that errors are generally concentrated in occlusion or small targets. Our method demonstrates excellent recognition rates for such complex targets, indicating the meaningful integration of SimAM attention mechanism and SPD-PAFPN in the RTMDet-Ins model. The experimental results show that our network is more suitable for high-precision segmentation of occluded *A. bisporus*.

Table 3. Segmentation results.

Sample ID	Number	Improved RTMDet-Ins		SOLOv2		CondInst		Watershed	
		Correct Number	Correct Rate (%)	Correct Number	Correct Rate (%)	Correct Number	Correct Rate (%)	Correct Number	Correct Rate (%)
No. 1	31	31	100.00	31	100.00	31	100.00	29	93.55
No. 2	41	41	100.00	39	95.12	41	100.00	38	92.68
No. 3	38	37	97.37	35	92.11	36	97.74	34	89.47
No. 4	33	33	100.00	33	100.00	33	100.00	30	90.91
No. 5	33	32	96.97	32	96.97	32	96.97	30	90.91
No. 6	44	43	97.73	41	93.18	41	93.18	39	88.64
No. 7	38	38	100.00	37	97.37	36	94.74	36	94.74
No. 8	29	29	100.00	28	96.55	28	96.55	28	96.55
No. 9	52	51	98.08	51	98.08	50	96.15	48	92.31
No. 10	53	51	96.23	50	94.34	51	96.23	49	92.45
Average Correct Rate		98.64%		96.37%		96.86%		92.22%	

3.5. Analysis of Center Positioning Experiment Results

The fitting effect of the location algorithm is shown in Figure 12, where the red box is the bounding box of the algorithm to the segmented target. Figure 12a represents the result of the segmentation algorithm, serving as the input to the localization algorithm. Figure 12b displays the final output of the localization algorithm, where the green ellipses demonstrate a high degree of fitting to the contours of the *A. bisporus*, thereby yielding precise picking point estimations marked by red dots. It can be demonstrated that the algorithm accurately fits and locates the *A. bisporus*.



(**a**) (**b**)

Figure 12. Localization results. (**a**) Origin image; (**b**) Fitting result.

Additionally, the experiment involved 15 samples of *A. bisporus*. After segmentation using the improved RTMDet-Ins algorithm, the manually annotated centers of *A. bisporus* were used as reference points for error evaluation. A comparison was made between the average errors of the least squares ellipse fitting localization algorithm and the traditional Hough circle fitting localization.

To more accurately evaluate the deviation of the calculated *A. bisporus* center, we introduced the Localization Error Rate (LER) [13] as the evaluation metric, expressed as follows:

$$\mathrm{LER} = \left(\left| \frac{c_m - c_a}{w} \right| + \left| \frac{r_m - r_a}{h} \right| \right) \times 100\% \tag{13}$$

where c_m and r_m represent the row and column coordinates of the manually determined center position of *A. bisporus*, and c_a and r_a represent the row and column coordinates of the algorithmically calculated center position. The values $w = 1856$ and $h = 640$ represent the width and height of the panoramic view containing the tested *A. bisporus*. The smaller the value of LER, the more accurate the positioning result. An example of LER calculation is shown in Table 4. The LER results of least square ellipse fitting and traditional Hough circle transform fitting are shown in Table 5.

Table 4. Example of LER calculation.

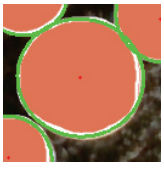
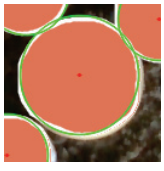
	Manual Positiomng	**Least-Squares Ellipse Fitting**	**Hough Transform Circle Fitting**
Positioning example			
Coordinate	(523, 235)	(523, 230)	(522, 228)
LER	0	0.78%	1.15%

Table 5. Localization results.

Sample ID	**Manual**	**Least-Squares Ellipse Fitting**		**Hough Transform Circle Fitting**	
		Algorithm Location	**LER (%)**	**Algorithm Location**	**LER (%)**
No. 1	(210, 71)	(211, 72)	0.21%	(208, 69)	0.42%
No. 2	(378, 261)	(380, 261)	0.11%	(380, 257)	0.73%
No. 3	(104, 338)	(102, 335)	0.58%	(102, 340)	0.42%
No. 4	(300, 420)	(303, 424)	0.79%	(303, 423)	0.63%
No. 5	(233, 178)	(235, 178)	0.11%	(233, 174)	0.63%
No. 6	(157, 409)	(152, 406)	0.74%	(155, 412)	0.58%
No. 7	(215, 295)	(218, 298)	0.63%	(213, 292)	0.58%
No. 8	(303, 255)	(308, 254)	0.43%	(306, 251)	0.79%
No. 9	(61, 183)	(61, 182)	0.16%	(57, 183)	0.22%
No. 10	(129, 133)	(129, 132)	0.16%	(128, 130)	0.52%
No. 11	(328, 163)	(329, 164)	0.21%	(325, 159)	0.79%
No. 12	(146, 243)	(146, 242)	0.16%	(148, 240)	0.58%
No. 13	(445, 384)	(446, 383)	0.21%	(441, 387)	0.68%
No. 14	(117, 44)	(118, 44)	0.05%	(118, 45)	0.21%
No. 15	(381, 29)	(380, 29)	0.05%	(379, 28)	0.26%
mean LER		0.31%		0.52%	

The experimental results show that the traditional Hough circle fitting localization algorithm successfully fitted all 15 samples, resulting in a mean LER of 0.52% after statistical analysis of the fitted results. In comparison, the ellipse fitting localization algorithm demonstrated superior performance. It also successfully delineated the contours of all *A. bisporus* samples, achieving a mean LER accuracy of 0.31%. The least squares ellipse fitting method reduced the absolute LER value by 0.21% compared to the Hough circle fitting algorithm. In summary, in complex scenarios, the least squares ellipse fitting method achieves higher accuracy in *A. bisporus* localization.

4. Conclusions

This study proposed a segmentation and localization algorithm for *A. bisporus* based on the improved RTMDet-Ins model. To enhance the algorithm's ability to extract features from small targets and address the challenge of accurately segmenting *A. bisporus* in complex environments, our method builds upon the RTMDet-Ins instance segmentation

network, incorporating the SimAM global attention mechanism and the lightweight feature fusion module SPD-PAFPN. The improved RTMDet-Ins algorithm reduces parameter count by 4.63% and decreases single-image computation time to 25.38 ms. In the *A. bisporus* test set, the AP^{50} reaches 98.50%, representing an improvement in detection accuracy and time efficiency compared to the original model. To further validate the algorithm's superiority, we compared our method with other classical segmentation networks using data collected by harvesting robots. Experimental results demonstrate that the improved RTMDet-Ins algorithm achieves a detection accuracy 98.64%, surpassing SOLOv2, CondInst, and traditional watershed algorithms. These findings indicate that the improved RTMDet-Ins instance segmentation algorithm can provide real-time, high-precision distribution information of *A. bisporus* for harvesting robots.

Secondly, a low-cost image stitching method based on the *wNCC* stitching algorithm with disparity correction is applied to generate high-precision panoramic multimodal fused images. The system was tested on 1320 photos to validate the model's effectiveness. The global stitching error is less than 2 mm within a 1200 mm $\times$ 400 mm. range. It provides high-precision panoramic information for subsequent algorithms.

The experimental results indicate that the method proposed in this paper exhibits superior segmentation and localization accuracy in complex harvesting environments. This can provide more precise and efficient visual perception for *A. bisporus* harvesting robots.

Although the method proposed in this paper has achieved satisfactory results in the scenarios of the collected dataset, differences in specific factory conditions, such as layer height and mushroom bed size, lead to significant deviations in captured images. This makes it challenging for directly transferred algorithms to achieve ideal results. Therefore, it is necessary to reacquire datasets and conduct annotation and training, incurring additional costs. Additionally, this design adopts an intermittent pause approach for image acquisition to ensure image capture quality, reducing the robot's overall operation efficiency. Therefore, future research will focus on improving the quality of images captured while in motion and enhancing the model's generalization ability, aiming to enhance efficiency further and reduce operational costs.

Author Contributions: Conceptualization, C.S. and C.Z. (Changsheng Zhu); methodology, C.S., C.Z. (Changsheng Zhu) and Y.M.; software, Y.M., X.R. and J.N.; validation, C.S., C.Z. (Changsheng Zhu), C.Z. (Chun Zhang), J.Y. and Y.M.; formal analysis, C.S., C.Z. (Chun Zhang), J.Y., Y.M. and X.R.; investigation, C.S., Y.M. and X.R.; resources, C.S. and J.Y.; data curation, Y.M., J.N. and X.R.; writing—original draft preparation, C.S., C.Z. (Changsheng Zhu), Y.M., J.N. and X.R.; writing—review and editing, C.S., C.Z. (Chun Zhang), C.Z. (Changsheng Zhu), Y.M., J.N. and J.Y.; visualization, Y.M. and J.N.; supervision, C.S., C.Z. (Chun Zhang), C.Z. (Changsheng Zhu) and J.Y. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by the Shandong Province Science and Technology SMEs Innovation Capability Enhancement Project (No. 2023TATSGC051 and No. 2023TATSGC014).

Institutional Review Board Statement: Not applicable.

Data Availability Statement: The data presented in this study are available on demand from the corresponding author at (zcs@sdust.edu.cn).

Conflicts of Interest: The author declares no conflicts of interest.

References

1. Sławińska, A.; Jabłońska-Ryś, E.; Gustaw, W. Physico-Chemical, Sensory, and Nutritional Properties of Shortbread Cookies Enriched with Agaricus bisporus and Pleurotus ostreatus Powders. *Appl. Sci.* **2024**, *14*, 1938. [CrossRef]
2. Iqbal, T.; Sohaib, M.; Iqbal, S.; Rehman, H. Exploring Therapeutic Potential of Pleurotus ostreatus and Agaricus bisporus Mushrooms against Hyperlipidemia and Oxidative Stress Using Animal Model. *Foods* **2024**, *13*, 709. [CrossRef] [PubMed]
3. Li, Y.; Pan, B.; Wan, Y. Research on the application and future development of visual recognition in modern agriculture. In Proceedings of the 2023 International Conference on Data Science, Advanced Algorithm and Intelligent Computing (DAI 2023); Atlantis Press: Amsterdam, The Netherlands, 2024; pp. 25–35.

4. Lu, C.P.; Liaw, J.J.; Wu, T.C.; Hung, T.F. Development of a mushroom growth measurement system applying deep learning for image recognition. *Agronomy* **2019**, *9*, 32. [CrossRef]
5. Wee, B.S.; Chin, C.S.; Sharma, A. Survey of Mushroom Harvesting Agricultural Robots and Systems Design. *IEEE Trans. AgriFood Electron.* **2024**, *2*, 59–80. [CrossRef]
6. Li, J.; Feng, Q.; Ru, M.; Sun, J.; Guo, X.; Zheng, W. Design of Shiitake Mushroom Robotic Picking Grasper: Considering Stipe Compressive Stress Relaxation. *Machines* **2024**, *12*, 241. [CrossRef]
7. Liu, J.; Liu, Z. The Vision-Based Target Recognition, Localization, and Control for Harvesting Robots: A Review. *Int. J. Precis. Eng. Manuf.* **2024**, *25*, 409–428. [CrossRef]
8. Eastwood, D.C.; Herman, B.; Noble, R.; Dobrovin-Pennington, A.; Sreenivasaprasad, S.; Burton, K.S. Environmental regulation of reproductive phase change in Agaricus bisporus by 1-octen-3-ol, temperature and CO_2. *Fungal Genet. Biol.* **2013**, *55*, 54–66. [CrossRef] [PubMed]
9. Reed, J.; Miles, S.; Butler, J.; Baldwin, M.; Noble, R. AE—Automation and emerging technologies: Automatic mushroom harvester development. *J. Agric. Eng. Res.* **2001**, *78*, 15–23. [CrossRef]
10. Yu, G.; Luo, J.; Zhao, Y. Region marking technique based on sequential scan and segmentation method of mushroom images. *Trans. Chin. Soc. Agric. Eng. (Trans. CSAE)* **2006**, *22*, 139–142.
11. Zhou, R.; Chang, Z.; Sun, Y.; Fan, P.; Tan, C. A Novel Watershed Image Segmentation Algorithm Based on Quantum Inspired Morphology. *J. Inf. Comput. Sci.* **2015**, *12*, 4331–4338. [CrossRef]
12. Chen, C.; Yi, S.; Mao, J.; Wang, F.; Zhang, B.; Du, F. A Novel Segmentation Recognition Algorithm of Agaricus bisporus Based on Morphology and Iterative Marker-Controlled Watershed Transform. *Agronomy* **2023**, *13*, 347. [CrossRef]
13. Yang, S.; Ni, B.; Du, W.; Yu, T. Research on an improved segmentation recognition algorithm of overlapping Agaricus bisporus. *Sensors* **2022**, *22*, 3946. [CrossRef] [PubMed]
14. Baisa, N.L.; Al-Diri, B. Mushrooms detection, localization and 3d pose estimation using rgb-d sensor for robotic-picking applications. *arXiv* **2022**, arXiv:2201.02837.
15. Shi, C.; Nie, J.; Mo, Y.; Zhang, C.; Zhu, C.; Zang, X. High precision scene stitching and recognition of agaricus bisporus based on depth camera. In Proceedings of the Third International Computing Imaging Conference (CITA 2023), Sydney, Australia, 1–3 June 2023; SPIE: Bellingham, WA, USA, 2023; Volume 12921, pp. 1001–1006.
16. Tian, Z.; Shen, C.; Chen, H. Conditional convolutions for instance segmentation. In Proceedings of the Computer Vision–ECCV 2020: 16th European Conference, Glasgow, UK, 23–28 August 2020; Proceedings, Part I 16; Springer: Berlin/Heidelberg, Germany, 2020; pp. 282–298.
17. Wang, X.; Zhang, R.; Kong, T.; Li, L.; Shen, C. Solov2: Dynamic and fast instance segmentation. *Adv. Neural Inf. Process. Syst.* **2020**, *33*, 17721–17732.
18. Zhong, M.; Han, R.; Liu, Y.; Huang, B.; Chai, X.; Liu, Y. Development, integration, and field evaluation of an autonomous Agaricus bisporus picking robot. *Comput. Electron. Agric.* **2024**, *220*, 108871. [CrossRef]
19. Briechle, K.; Hanebeck, U.D. Template matching using fast normalized cross correlation. In Proceedings of the Optical Pattern Recognition XII, Orlando, FL, USA, 16–20 April 2001; SPIE: Bellingham, WA, USA, 2001; Volume 4387, pp. 95–102.
20. Lyu, C.; Zhang, W.; Huang, H.; Zhou, Y.; Wang, Y.; Liu, Y.; Zhang, S.; Chen, K. Rtmdet: An empirical study of designing real-time object detectors. *arXiv* **2022**, arXiv:2212.07784.
21. Yang, L.; Zhang, R.Y.; Li, L.; Xie, X. Simam: A simple, parameter-free attention module for convolutional neural networks. In Proceedings of the International Conference on Machine Learning, Virtual, 18–24 July 2021; PMLR: New York, NY, USA, 2021; pp. 11863–11874.
22. Chollet, F. Xception: Deep learning with depthwise separable convolutions. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Honolulu, HI, USA, 21–16 July 2017; pp. 1251–1258.
23. Lin, T.Y.; Dollár, P.; Girshick, R.; He, K.; Hariharan, B.; Belongie, S. Feature pyramid networks for object detection. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Honolulu, HI, USA, 21–16 July 2017; pp. 2117–2125.
24. Liu, S.; Qi, L.; Qin, H.; Shi, J.; Jia, J. Path aggregation network for instance segmentation. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Salt Lake City, UT, USA, 18–23 June 2018; pp. 8759–8768.
25. Sunkara, R.; Luo, T. No more strided convolutions or pooling: A new CNN building block for low-resolution images and small objects. In Proceedings of the Joint European Conference on Machine Learning and Knowledge Discovery in Databases, Grenoble, France, 19–23 September 2022; Springer: Berlin/Heidelberg, Germany, 2022; pp. 443–459.
26. Fausto Milletari, N.; V-Net, A.S.A. Fully Convolutional Neural Networks for Volumetric Medical Image Segmentation. Available online: https://arxiv.org/abs/1606.04797 (accessed on 2 January 2024).
27. Wang, C.Y.; Liao, H.Y.M.; Wu, Y.H.; Chen, P.Y.; Hsieh, J.W.; Yeh, I.H. 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, Seattle, WA, USA, 14–19 June 2020; pp. 390–391.

28. Sajjadi, M.S.; Vemulapalli, R.; Brown, M. Frame-recurrent video super-resolution. In Proceedings of the IEEE Conference on Computer Vision and pattern Recognition, Salt Lake City, UT, USA, 18–23 June 2018; pp. 6626–6634.
29. Fitzgibbon, A.; Pilu, M.; Fisher, R.B. Direct least square fitting of ellipses. *IEEE Trans. Pattern Anal. Mach. Intell.* **1999**, *21*, 476–480. [CrossRef]

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.

agriculture

MDPI

Article

Super-Resolution Semantic Segmentation of Droplet Deposition Image for Low-Cost Spraying Measurement

Jian Liu [1], Shihui Yu [1], Xuemei Liu [1,2,*], Guohang Lu [1], Zhenbo Xin [1,2] and Jin Yuan [1,2]

[1] College of Mechanical & Electronic Engineering, Shandong Agricultural University, Tai'an 271018, China; unknownlogic@163.com (J.L.); yushihuilearn@163.com (S.Y.); gh.lu_701@sdau.edu.cn (G.L.); xinzb@sdau.edu.cn (Z.X.); jyuan@sdau.edu.cn (J.Y.)
[2] Shandong Agricultural Equipment Intelligent Engineering Laboratory, Tai'an 271018, China
* Correspondence: lxmywj@126.com; Tel.: +86-152-6986-8559

Abstract: In-field in situ droplet deposition digitization is beneficial for obtaining feedback on spraying performance and precise spray control, the cost-effectiveness of the measurement system is crucial to its scalable application. However, the limitations of camera performance in low-cost imaging systems, coupled with dense spray droplets and a complex imaging environment, result in blurred and low-resolution images of the deposited droplets, which creates challenges in obtaining accurate measurements. This paper proposes a Droplet Super-Resolution Semantic Segmentation (DSRSS) model and a Multi-Adhesion Concave Segmentation (MACS) algorithm to address the accurate segmentation problem in low-quality droplet deposition images, and achieve a precise and efficient multi-parameter measurement of droplet deposition. Firstly, a droplet deposition image dataset (DDID) is constructed by capturing high-definition droplet images and using image reconstruction methods. Then, a lightweight DSRSS model combined with anti-blurring and super-resolution semantic segmentation is proposed to achieve semantic segmentation of deposited droplets and super-resolution reconstruction of segmentation masks. The weighted IoU (WIoU) loss function is used to improve the segmented independence of droplets, and a comprehensive evaluation criterion containing six sub-items is used for parameter optimization. Finally, the MACS algorithm continues to segment the remained adhesive droplets processed by the DSRSS model and corrects the bias of the individual droplet regions by regression. The experiments show that when the two weight parameters α and β in WIoU are 0.775 and 0.225, respectively, the droplet segmentation independence rate of DSRSS on the DDID reaches 0.998, and the IoU reaches 0.973. The MACS algorithm reduces the droplet adhesion rate in images with a coverage rate of more than 30% by 15.7%, and the correction function reduces the coverage error of model segmentation by 3.54%. The parameters of the DSRSS model are less than 1 M, making it possible to run it on embedded platforms. The proposed approach improves the accuracy of spray measurement using low-quality droplet deposition image and will help to scale-up of fast spray measurements in the field.

Keywords: droplet deposition measurement; image processing; super-resolution; semantic segmentation

Citation: Liu, J.; Yu, S.; Liu, X.; Lu, G.; Xin, Z.; Yuan, J. Super-Resolution Semantic Segmentation of Droplet Deposition Image for Low-Cost Spraying Measurement. *Agriculture* **2024**, *14*, 106. https://doi.org/10.3390/agriculture14010106

Academic Editor: Jianli Song

Received: 4 December 2023
Revised: 29 December 2023
Accepted: 2 January 2024
Published: 8 January 2024



Copyright: © 2024 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (https://creativecommons.org/licenses/by/4.0/).

1. Introduction

Precision agriculture, arising from the need to mitigate the environmental impact and enhance agricultural input efficiency, necessitates the meticulous control of pesticide application. This places significant demands on the spraying quality of land sprayers and drones. In order to carry out high-quality spraying operations, on the one hand, new types of sprayers with higher performance have been developed [1], and on the other hand, it is necessary to monitor the quality of the spraying in real time, in particular to monitor droplet deposition on the leaves of the plant in a comprehensive manner. The monitoring measurements can then be utilized as feedback data to support the adjustment of the sprayer's spraying parameters accordingly, which serves as the key to reducing pesticide

application and improving overall efficiency. To ensure the real-time and convenient monitoring of plant droplet deposition, it becomes essential to enhance the automatic digitization of measuring droplet deposition on plant leaves. Furthermore, it is crucial that we reduce the time delay associated with individual measurements and develop rapid in situ measurement techniques. These endeavors have emerged as a focal point of research in this field [2–6].

Indicators for measuring spray quality include droplet coverage, droplet density, Volume Median Diameter (VMD), Number Median Diameter (NMD), etc. Calculating these indicators requires counting the number of droplets and the volume or diameter of each individual droplet. However, in recent years, as drones and sprayers have widely used precise nozzles for high-quality pesticide application operations, the droplets from agricultural sprayers are generally relatively dense, and the evaluation needs to be completed with the help of sophisticated test strips or devices. Currently, the spray measurement method with the widest range of applications and the most mature technology is the water-sensitive paper method (WSP). The process of implementing this method includes steps such as placement, spraying, recycling, scanning, and analyzing. In the early application of the WSP method, scanning and analyzing needed to be performed in the laboratory, and transferring the water-sensitive paper consumed a lot of time. Zhu et al. used a portable business card scanner, a portable computer, and a customized software package called DepositScan version 1.2 to form a system which digitized the method to a certain extent and greatly improved the speed [7]. However, the above method still relies on the manual operation of WSP, such as placing and retrieving and discrete digitalization, and the consumption of WSP still restricts the sustainability of this method. The consumption of water-sensitive paper during large-scale droplet measurement brings excessive economic costs.

Recently, machine vision has been utilized in spray measurement to acquire the details of droplets deposited by agricultural sprays. Wang et al. designed a new piece of machine vision-based droplet collection equipment, using recyclable and erasable oil paper instead of WSP [4]. The device is less disturbed by the environment and can realize automatic control and the automation of spray deposition image shooting. The collected images are segmented by a watershed algorithm to achieve droplet segmentation. Based on the above equipment, Wang, et al. proposed a droplet detection model based on SSD_MobileNet, which reconstructed the droplet outline based on the detection results, further improving the statistical accuracy of the droplet parameters [8]. Yang, et al. used the deep learning semantic segmentation model Deeplab V3+ to realize the segmentation of deposited droplets, ensuring the accuracy of droplet segmentation [6]. This study considers clusters of adherent droplets and independent droplets as two different object categories. The neural network is used to find the concave points in the adherent droplets, and the concave segmentation algorithm is used to individually separate the adherent droplet clusters. The rate of recognizing adhesion pits reaches 95%, which can effectively identify and segment most of the adhered droplets, thus effectively improving the accuracy of the spray measurement.

When droplets are deposited on WSP and erasable oil paper, they are in an absorption and diffusion state, which is completely different from the form of droplets deposited on leaves. The area which is stained by the deposition of spray on WSP and erasable oil paper is larger than the area covered on the leaf surface. Under a high-intensity spray, it is easy for the overlap of stained patches to occur, affecting the measurement results. Using a non-water-absorbent material as the surface for the deposition of the droplets can simulate the hemispherical state of droplets when deposited on leaves. Vacalebre, et al. studied the wetting characteristics of acrylic material surfaces. It was found that the acrylic surface treated with a specific method can obtain strong hydrophobicity and a stable WCA (wetting angle) [9]. The volume of a droplet can be calculated from its deposition area using the known wetting agent angle [10]. According to the above principle, an Optical Droplet Edge Imaging (ODEI)-based automatic spray measurement method has been designed. Its measurement accuracy is directly related to the performance of the

camera equipped with the measurement device, but high-performance cameras are not conducive to their promotion in scalable application. However, the limitations caused by camera performance in a low-cost imaging system, coupled with dense spray droplets and a complex imaging environment, result in blurred and low-resolution images of deposited droplets, which makes obtaining accurate measurements challenging. How to achieve precise measurements by leveraging low-cost cameras becomes a practical research topic.

In recent years, deep learning has made great progress in the fields of image super-resolution reconstruction and deblurring. Image deblurring models such as DeblurGAN [11,12] can repair image blur, and its effect is better than traditional methods such as Wiener filtering and deconvolution; Quan, et al. proposed using a linear mixture of a series of Gaussian blur kernels to approximately express out-of-focus blur, and built the end-to-end deep learning model GKMNet to correct de-focus blur with remarkable results [13]. Super-resolution reconstruction models such as SRGAN [14] can reconstruct image details when enlarging the image, and its effect is far better than traditional methods based on interpolation. The deep learning segmentation model can consider the semantic information of pixels when segmenting images and identify light spots as being parts of droplets, which is better than traditional methods. Some research combines super-resolution ideas with semantic segmentation methods to improve segmentation accuracy [15].

This paper reconsiders the method of processing images of droplet deposition acquired by low-cost cameras, merging deblurring, super-resolution, and image segmentation to build an end-to-end spray image processor that meets the requirements of rapid in situ spray measurement for scalable application. The specific contributions of this paper are as follows:

1. A droplet image semantic segmentation model is built to output a super-resolution segmentation mask to accurately express the deposited area of individual droplets.
2. The weighted IoU loss function is applied to improve the statistical independence of droplets in super-resolution semantic segmentation.
3. The concave segmentation method is applied for complex droplet adhesion clusters segmentation, including annular type and multiple adhesion.

2. Materials and Methods

2.1. Characteristics of Deposited Droplet Images

A portable in situ droplet deposition measurement device was developed to validate the ODEI method. The process of collecting droplet deposition images in the crop canopy using the in situ measurement device is shown in Figure 1. The user controls the acrylic collector to eject to collect spraying droplets with the embedded system, then retracts the collector to capture images. The image processing program embedded in the device analyzes the image and measures the parameters of droplets. In order to measure the deposition of droplets on both sides of the crops' leaves, an LED surface light source is placed above both sides of the collector to obliquely illuminate the front and back sides of the collector where the droplets have been collected. The reflector reflects the images of the front and back sides of the collector to the camera to record a digital image. The optical principles of the device are shown in Figure 2. In practical applications, if the device can be equipped with a camera with lower specifications, the application cost of the method can be reduced, which is conducive to the promotion of scalable application. However, the limitations caused by camera performance in a low-cost imaging system, coupled with a dense spray of droplets and a complex imaging environment, result in blurred and low-resolution images of deposited droplets, which challenges the accurate measurements. The specifical characteristics of low-quality images are as follows:

(1) The fog droplet deposition image obtained shows a phenomenon of blurry edges and clear middle.
(2) The blurring at the edge of the image leads to the dissolution of the droplet boundaries in the edge area, and adjacent non adhesive droplets are prone to displaying adhesive properties in the blurred image.

(3) The relatively low resolution cannot accurately show the contour of the edge small droplets, resulting in insufficient accuracy in the statistics for the droplet deposition area.
(4) The ODEI method forms high brightness spots in the center of the deposited droplets, which makes droplet segmentation more difficult.
(5) The surface of the collector shows dirt and scratches after many measurements, resulting in a decrease in the quality of the images collected.

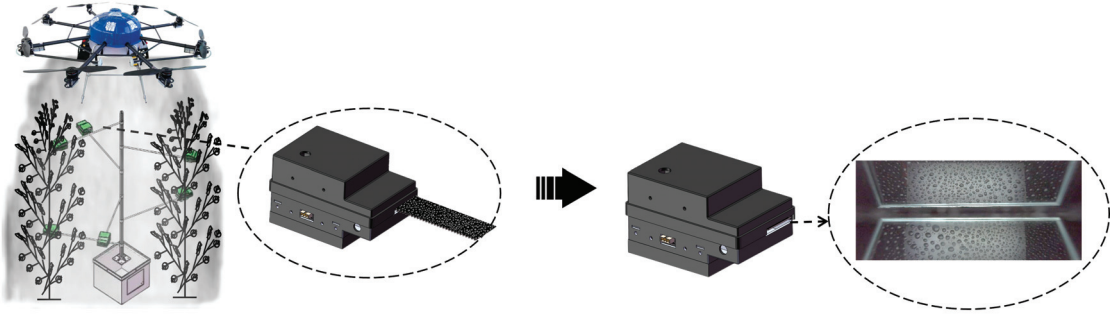


Figure 1. The process of collecting droplet deposition images using in situ measurement devices.

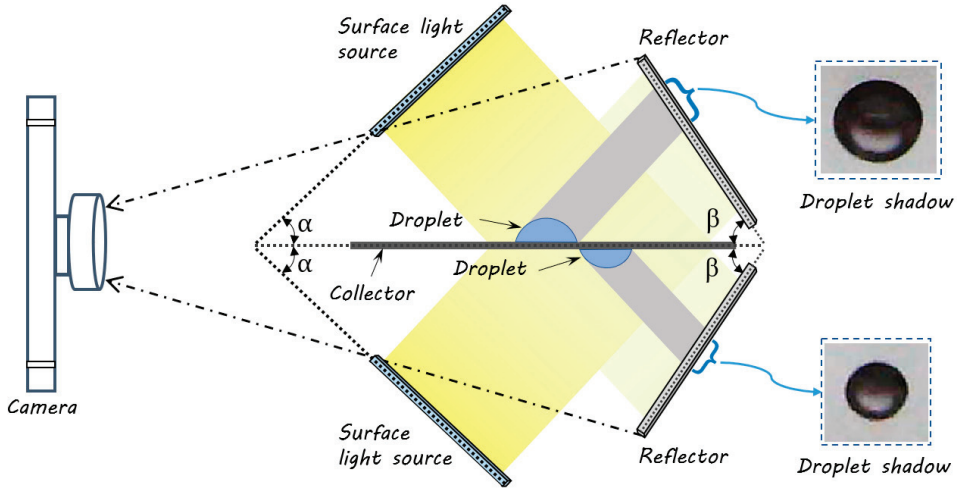


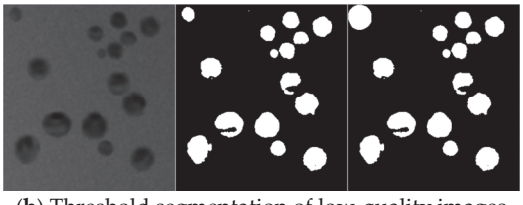
Figure 2. Optical principle of image acquisition device for spray deposition.

The key to spray deposition image processing is to identify individual droplets, calculate their deposition area, convert the deposition area into volume, and then calculate VMD, NMD, and other droplet parameters. As in the device shown in Figure 2, under the illumination of the white LED surface light source, the droplets form an image in the camera as a black ring with a white spot in the center, and the non-droplet area forms an image in the camera with a white background. Calculating the area of individual droplets requires segmenting them from the digital image. Under ideal conditions, the white spot is wrapped by the shadowed edge of the droplet. Although the color and brightness of the white spot and the white background are very close, segmented individual droplets can be achieved by filling them after threshold segmentation, as shown in Figure 3a. However, limited by the performance bottleneck of the camera and some environmental factors, the collected images may be blurry and have insufficient resolution. Severe blur may cause the white spot to deviate from the droplet edge, causing the threshold segmentation and filling method to be ineffective, as shown in Figure 3b. Therefore, accurately obtaining the

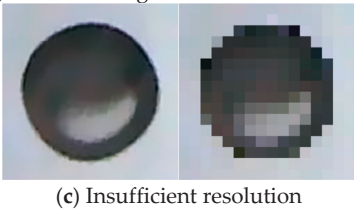
area of deposited individual droplets in low-quality deposition images is a challenging task. Insufficient resolution can also lead to inaccurate statistical results for the droplet deposition area. If the number of pixels constituting the droplet is too small, it is hard to accurately characterize the deposition boundary of the droplet, as shown in Figure 3c. Droplet adhesion is another problem that makes droplet area statistics difficult. If the droplets that successively fall on the collector come into physical contact, they will merge into a single droplet and form an elliptical deposition surface under the combined action of surface tension and adhesion. If two droplets are not in physical contact but are very close, image blurring may lead to the edges of the droplets dissolving (blur, spread) and invading each other, forming a connected domain in the image, as shown in Figure 3d, ultimately making it difficult to segment individual droplets. In summary, processing low-quality deposition images is a complex task, and droplet segmentation algorithms need to have the ability to identify individual droplets.



(**a**) Threshold segmentation in ideal state



(**b**) Threshold segmentation of low-quality images



(**c**) Insufficient resolution



(**d**) Imaging blurring causing adhesions

Figure 3. Threshold based droplet segmentation in ideal and actual situations.

2.2. Workflow of Droplet Deposition Low-Cost Measurement

The method proposed in this study aims to construct an image processing process to achieve the measurement of parameters based on low-quality droplet deposition images. Super-resolution semantic segmentation based on deep learning is the core of this method. Datasets, model structures, loss functions, and droplet segmentation algorithms are improved to improve the method's processing capabilities for spray deposition images. The workflow of the method is shown in Figure 4. The blue and green arrows in the figure indicate two processes: the training process and the working process.

Training stage: (1) In order to train the model, image reconstruction-based data augmentation methods were used to create the DDID. The samples included use low-quality images as model input and high-resolution segmentation masks as segmentation annotations. (2) In order to solve the blur problems, insufficient resolution, and imaging adhesion in low-quality droplet deposition image segmentation, a lightweight super-resolution semantic segmentation model DSRSS was constructed to output super-resolution segmentation masks to represent the size and boundaries of individual droplets. (3) In order to improve the droplet independence rate in the DSRSS's output results, a weighted IoU loss function was applied to optimize the category weights, finally achieving high-quality droplet segmentation.

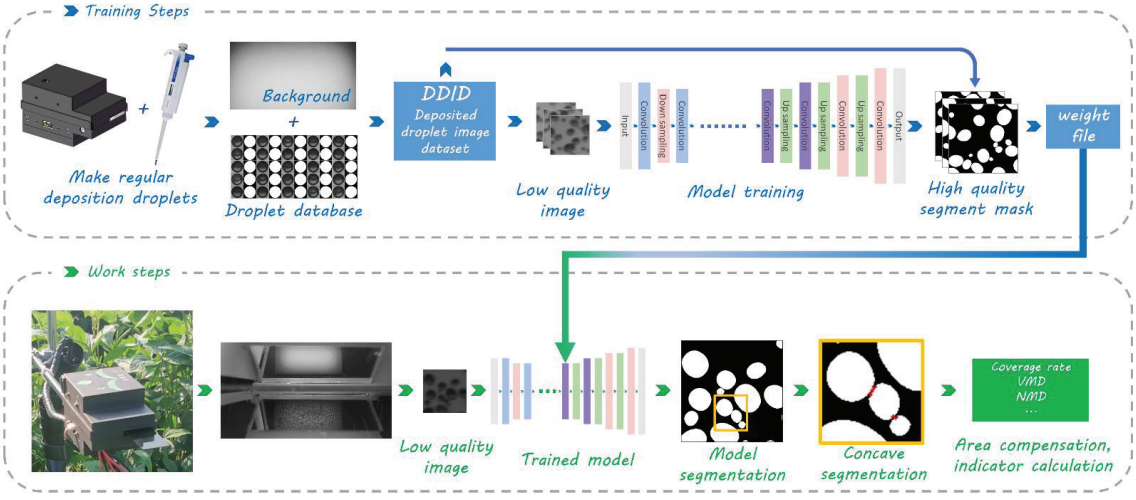


Figure 4. The workflow of parameter acquisition in low-cost droplet deposition measurement.

Working stage: (1) The device collects droplet deposition images and performs preprocessing, including image cropping and distortion correction. (2) The preprocessed image was input into the DSRSS, and through the model's inference computation, a super-resolution segmentation mask end-to-end model was output. (3) The super-resolution segmentation mask is processed with a concave point segmentation algorithm to further improve the independence of droplet segmentation. (4) Droplet measurement indicators, such as coverage rate, droplet density, and VMD, are calculated from the segmentation mask after concave point segmentation.

2.3. Construction of Deposition Droplet Image Dataset

2.3.1. Feasibility of Droplet Deposition Image Generation

In order for the model to learn richer morphological patterns of deposited droplets, the annotated GT (ground truth) of the sample must represent the deposition surface of all droplets as accurately as possible, which requires the high-quality image used to annotate the segmentation mask to have a sufficiently high resolution and clarity. An ideal dataset creation method is to use a high-performance camera and a low-performance camera to capture images at the same location. Segmentation masks are annotated with high-quality images taken by high-performance cameras, and they are combined with images taken by low-standard cameras to form training sample pairs. However, it is difficult to capture images with a highly consistent field of view because two cameras with different specifications will generate pose errors, perspective errors, and distortion correction errors during the shooting process, not to mention that the water droplets will evaporate and move during the image switching process. It is difficult to meet these requirements with directly captured images.

According to capillary and wetting phenomena [10], when a droplet with a diameter smaller than the capillary length of the liquid is deposited on a flat solid surface, its water contact angle is constant and appears as a spherical crown with a consistent shape. That is to say, within a reasonable particle size range, taking high-definition images of large droplets and reducing their resolution will result in images that will be highly consistent with directly captured images of small droplets. Therefore, it is feasible to use the data augmentation method of image reconstruction to construct a training data set.

2.3.2. High Quality Image Samples Generation

A large droplet shape dataset containing 67,829 samples was established as the candidate units for generating image samples by extracting individual droplets from high-quality deposition images captured with a high-specification camera. The droplet morphology database includes the grayscale magnification of the droplet region pixels relative to the background pixels in the region and the local segmentation mask of the droplet individual, as shown in Figure 5. The original position of droplets on the collector were also recorded in the database to preserve any correlation with droplet morphology and position. Individual droplets in these databases were used as the foreground during sample generation, with the background coming from empty collector images captured using the ODEI device. During the sample image generation process, individual droplets are scaled to random sizes and attached to the background image according to their original positions recorded in the database, without overlapping each other. Considering that the droplets tilted onto the collector and the droplets which were fused due to physical contact deposit in a long elliptical shape, random shear transformation was also performed on individual droplets. At the same time, the same transformation is applied to the corresponding segmentation mask, and the resulting image sample pair is shown in Figure 6.

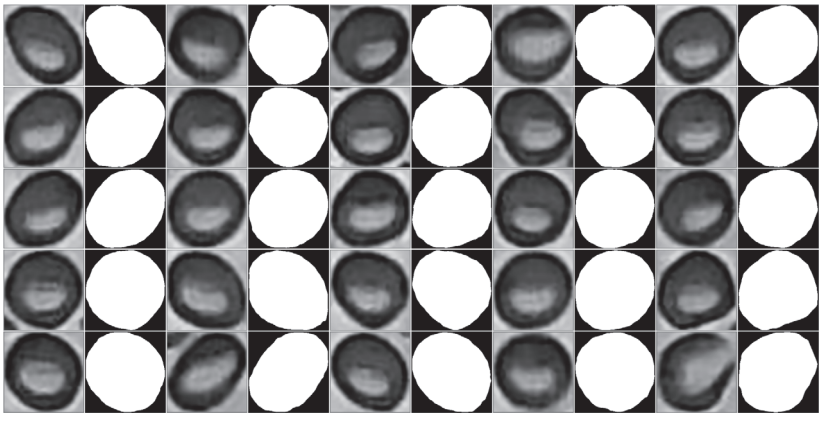
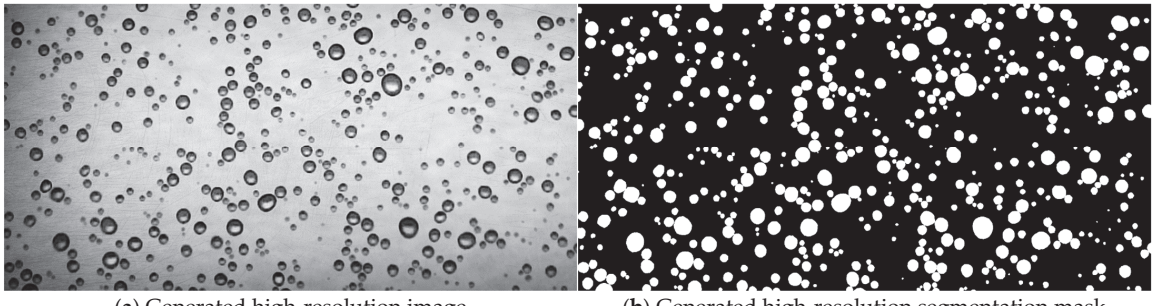


Figure 5. Partial samples from the droplet morphology database.



(**a**) Generated high-resolution image (**b**) Generated high-resolution segmentation mask

Figure 6. High quality image and synchronously generated segmentation annotations.

2.3.3. Quality-Degrading Operations for the Simulation of Low-Quality Image Acquisition

Low image quality is the result of multiple factors, the most important factors are insufficient COMS sensor performance, which leads to detail loss, and low-performance ultra-wide-angle lenses, which cause spherical aberration. The low-quality images in the dataset are generated from high-quality images through a series of operations, including reducing resolution, adding noise, and applying various fuzzy transformations. Among them,

reducing the resolution is used to simulate the insufficient resolution of low-performance cameras, image noise (including salt and pepper noise and Gaussian noise) and Gaussian blur are used to simulate the loss of image details, and radial blur which gradually increases from the center to the edge of the image is used to simulate spherical aberration, resulting in image margin blurring. Therefore, stains and scratch textures were randomly added to the low-quality images to simulate real images collected by the device in field applications.

Since scratches are physically present textures on the collector, the scratch simulation in the generated image is directly performed on the high-resolution image background, prior to the attachment of droplets. The scratch simulation has actually been completed in the generated image shown in Figure 6a. Spherical aberration is caused by the abnormal refraction of light by the lens, which occurs before the imaging of CMOS sensors, so the application of radial blur is immediately followed by scratch simulation. Gaussian blur, adding noise, and reducing resolution are all implemented to simulate image detail loss. However, Gaussian blur reduces the effect of salt and pepper noise, and reducing resolution weakens the effect of Gaussian blur. Therefore, the order of executing these three steps is: reducing resolution, Gaussian blur, and adding noise. The simulation process for the low-quality images is shown in Figure 7, and a comparison between the generated low-resolution images and real low-quality images is shown in Figure 8.

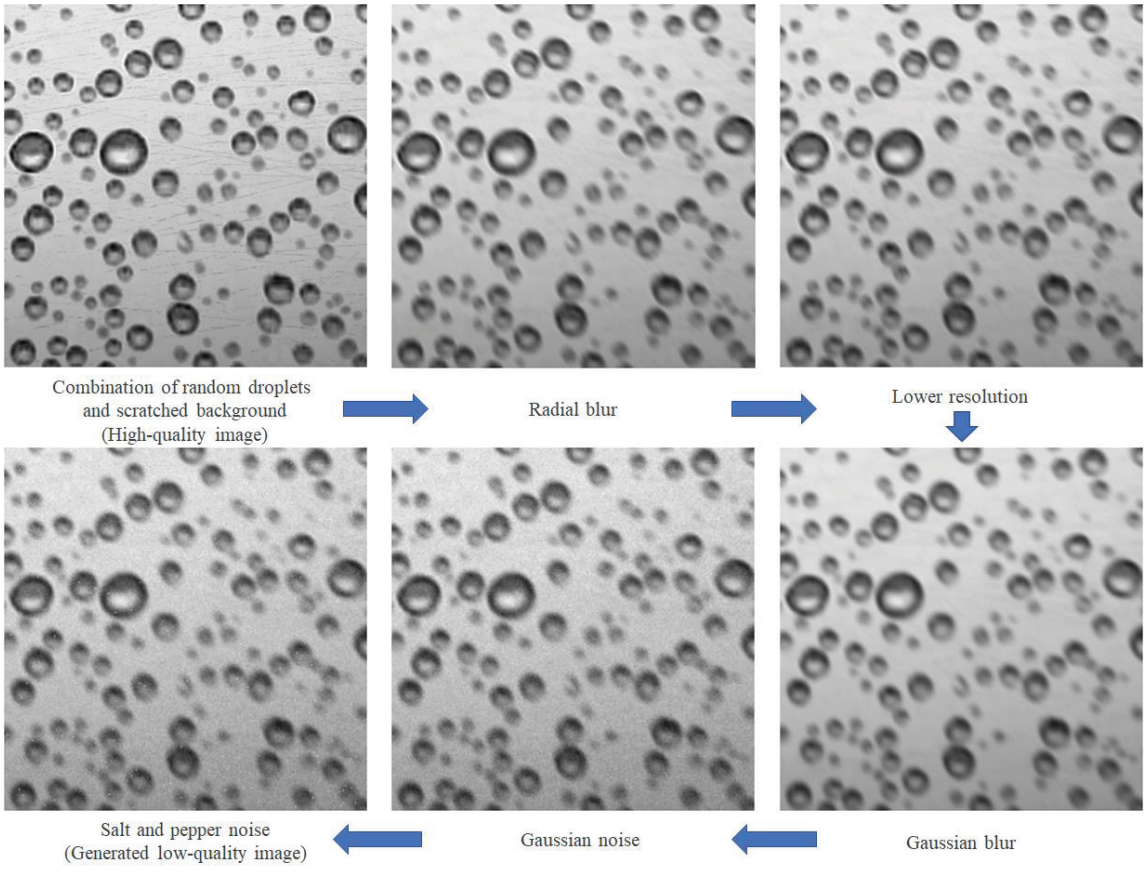
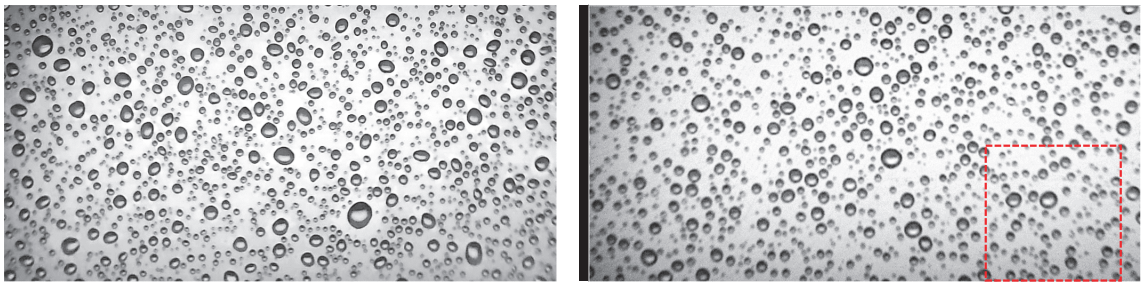


Figure 7. The simulation process for low-quality images.



(**a**) Real low-quality image (**b**) Generated simulated low-quality image

Figure 8. Low quality image samples. (The part in the dotted box is the area shown in Figure 7).

In order to meet the needs of model training, validation, and comparison, each group of samples in the dataset also contains additional sub items, which come from the intermediate process of dataset generation. Finally, a total of 2000 groups of training samples were generated, of which 1600 groups were used as the training set and 400 groups were used as the validation set. In this paper, the above dataset was named DDID (deposited droplet image dataset).

2.4. Droplet Super-Resolution Semantic Segmentation (DSRSS) Method

2.4.1. Structure of the DSRSS

In the field of deep learning image segmentation, researchers use large-scale natural image datasets to verify the versatility and generalization ability of the model. The scale of SOTA models is getting larger and larger, including an increasing number of special structures, which also puts forward increasing requirements for computer performance. Compared with natural image datasets, the information contained in spray deposition images is relatively simple. Applying large models to solve the segmentation of droplet images will cause computer hardware resources to be wasted and is not conducive to cost control. Due to the above reasons, this study built a lightweight CNN model to meet the requirements for application in embedded systems.

The idea of super-resolution was used to solve the problem of insufficient resolution of small droplets. DSRSS is a super-resolution semantic segmentation model used for dealing with low-quality droplet deposition images. The length and width of the output segmentation mask are four times that of the input image, which can more accurately express the morphology of the outline of the droplets. In common super-resolution image reconstruction tasks [14,16], the deep learning model learns the prior knowledge contained in large-scale datasets during training and applies it to the restoration of image details such as color and texture, which places higher requirements on the number of parameters in the model. The super-resolution semantic segmentation task does not consider complex image details and only focuses on outputting more accurate segmentation masks. It has lower requirements for the number of model parameters and is conducive to lightweighting the model.

DSRSS adopts an encoder–decoder structure, which is widely used in semantic segmentation models, such as U-Net [17] and its derivative models, SegNet [18], DeepLab V3+ [19], etc. The function of the encoder is feature extraction, and the decoder generates a segmentation mask based on the features. Under the guidance of the training data, the decoder can learn some prior knowledge and act on image reconstruction.

DSRSS is a fully convolutional network that does not contain fully connected layers [20]. It uses standard 2D convolution operations as convolutional layers, and batch normalization and swish [21] function activation are performed after each convolutional layer. On the encoder side, the activation function is followed by a pooling operation with a kernel size of 2×2, and on the decoder side; a bilinear interpolation up-sampling operation with a factor of 2×2 is performed before each convolutional layer. The cross-layer connection

structure in U-Net was referenced. The feature map obtained by each down-sampling in the encoder is concatenated with the feature map of the same level in the decoder. In the test, this structure can significantly improve the *IoU* of the model output.

The structure of the DSRSS model has been determined through a large number of experiments, with its depth and width at relatively optimal values. Cross layer connection structure is also essential. The down-sampling level must reach at least 3 to ensure the high performance of the model. If the down-sampling level is lower than 3, errors in the segmentation of droplet spots are prone to occur. To achieve 4× super-resolution output, the number of up-sampling operations should be 2 more than the number of pooling operations. The width of the model also has a great impact on the performance of the model, but the larger the scale of the model, the higher the requirements for hardware performance. The final model structure adopted is shown in Figure 9. On this basis, the benefits of continuing to increase the width of the model are no longer clear.

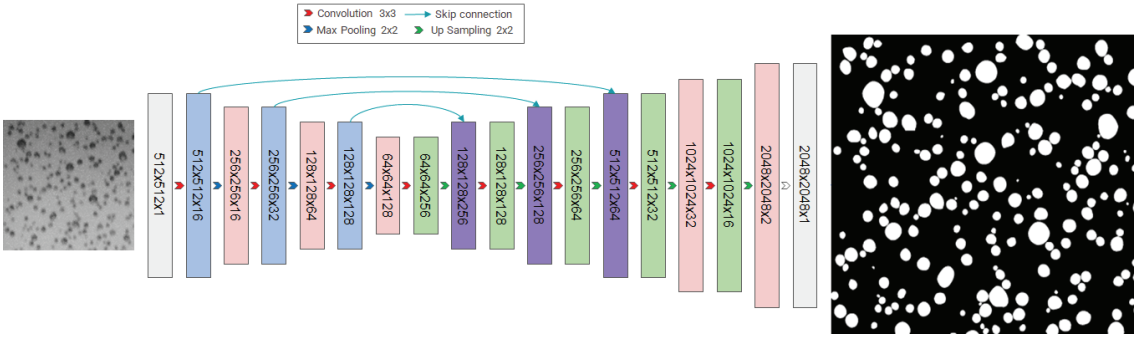


Figure 9. Structure of DSRSS.

2.4.2. Loss Function and Metrics

The independence of individual droplets and the accuracy of segmentation areas are the two most important criteria in measuring the quality of DSRSS segmentation, and the loss function determines the model's optimization goal. Semantic segmentation does not pay attention to the independence of individual droplets, and its optimization is usually directed towards improve the accuracy of segmented regions. In order to enable the model to take into account the independence of individual droplets, specific metrics and loss functions were applied.

Three normalization metrics, including the recognition rate, independence rate, and correction rate, of individual droplets were used to characterize the independence of individual droplets. The calculation method is: (1) Search for contours in the segmentation mask output by the model to obtain a list of predicted droplets. (2) Search for contours in GT to obtain a list of target droplets. (3) Traverse the target droplet list, count the proportion of individuals that only intersect with one predicted droplet and this intersecting predicted droplet has no other intersecting droplets as the independence rate, and count the proportion of individuals that intersect with any predicted droplet as the recognition rate. (4) Traverse all predicted droplets and count the proportion of individuals that intersect with any target droplet as the accuracy rate. The three quantity standards can be expressed in Equation (1).

$$\begin{cases} T_r = \dfrac{N_{true}}{N_t} \\ I_r = \dfrac{N_{indep}}{N_t} \\ C_r = \dfrac{N_{true}}{N_p} \end{cases} \tag{1}$$

where, T_r denotes the recognition rate, I_r denotes the independence rate, C_r denotes the correct rate, N_t denotes the number of target droplets, N_p denotes the number of predicted droplets, N_{true} denotes the number of correctly recognized droplets, and N_{indep} denotes the number of droplets without adhesion.

The *IoU* is often used as one of the evaluation criteria for semantic segmentation and is also used to define the loss function. The *IoU* loss function optimizes the intersection area and union area ratio of the target area and the prediction area toward 1, so that the segmentation map output by the model is as consistent as possible with GT. The *IoU* loss function and its difference set form a two-category segmentation which can be expressed by Equation (2).

$$L_{IoU} = 1 - \frac{A_t \cap A_p}{A_t \cup A_p} = \frac{(A_t \backslash A_p) + (A_p \backslash A_t)}{A_t \cup A_p} \tag{2}$$

where A_t denotes the target area, A_p denotes the prediction area, and L_{IoU} denotes the loss value.

In the segmentation mask output by the model, the droplet adhesion points are composed of very few pixels. However, the segmentation of these pixels determines the model's ability to segment individual droplets, so their segmentation should be of higher priority. This special priority can be reflected in the loss function. Since the adhesion points between droplets are only included in the difference set $A_t \backslash A_p$ between the prediction area and the target area, reducing the proportion of this difference set in the loss function can adjust the sensitivity of the model to droplet adhesions. So, two weight values α and β were added into Equation (2), and the loss value is represented by the symbol L_{IoU}. The modified loss function was named *WIoU*, as shown in Equation (3).

$$L_{WIoU} = \frac{\alpha * (A_t \backslash A_p) + \beta * (A_p \backslash A_t)}{A_t \cup A_p} \tag{3}$$

The imbalance of weights α and β will change the optimization direction of the model, improve the independence rate, and lead to a decrease in *IoU* metric. Precision and Recall can reflect the above influence. These two evaluation indicators can be represented with Equation (4), and the meaning of the symbols is the same as Equation (1).

$$\begin{cases} Precision = \dfrac{A_t \cap A_p}{A_p} \\ Recall = \dfrac{A_t \cap A_p}{A_t} \end{cases} \tag{4}$$

The *IoU* can be regarded as a combination of *Precision* and *Recall*. Theoretically, the larger the weight α and the smaller the weight β in Equation (3), the higher the Precision of the model and the lower the probability of adhesion; however, the Recall will decrease. Therefore, there must be an optimal combination of α and β that can balance the independence of the droplets and the accuracy of the segmentation area to achieve the optimal segmentation effect.

The sum of a total of six indicators including *IoU*, *Precision*, and *Recall*, which characterize the accuracy of the region, and T_r, I_r, and C_r, which characterize the independence of the droplets, represented by the symbol M, was used as the criterion for judging the quality of the model segmentation effect to estimate the optimal weight coefficients α and β, as shown in Equation (5).

$$M = IoU + Precision + Recall + T_r + I_r + C_r \tag{5}$$

2.5. Multi-Adhesion Concave Segmentation (MACS) Algorithm

The output of the DSRSS model does not contain individual information, and additional algorithms are required to extract individual droplets and calculate their parameters.

Although the droplet independence rate in the DSRSS model's output is already very high after WIoU loss function optimization, it cannot reach 100%, especially when processing extremely low-quality images. Therefore, when extracting individual droplets, it is necessary to further segment the adherent droplets.

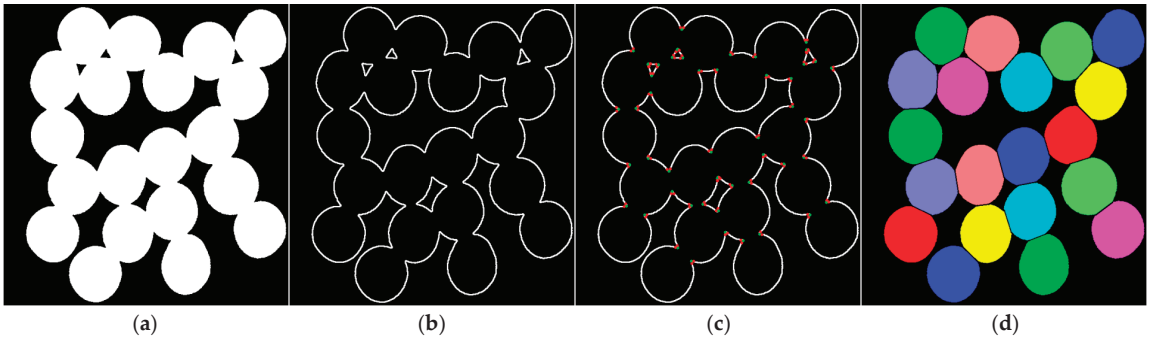
Contours were used to represent individual droplets, and the roundness of the contour is used to determine whether there is adhesion in the extracted droplet patches. Contours with a roundness greater than or equal to the threshold are judged as independent droplets, otherwise they are judged as adherent droplets. The formula for calculating roundness is shown in Equation (6).

$$R = \frac{4\pi A}{L^2}, 0 < R < 1 \tag{6}$$

where A denotes the area of the segmented patch, L denotes the perimeter of the segmented patch, and R denotes the roundness of the segmented patch.

The algorithm for separating adherent droplets is the concave point method, which has many implementation methods. The most critical step in the concave point method is to locate the concave points in the image. In the high-resolution segmentation mask output by the DSRSS model, the adhesion point of the droplet has a sharp contact angle, so the concave point of the adhesion point can be regarded as a corner point and located through the contour corner detection algorithm. Another key step in the concave segmentation method is the pairwise matching of concave points. Since the deposition bottom surfaces of droplets are all approximately elliptical, the slope of a straight line passing through a pair of concave points must be between the slope of the straight line connecting the deepest point of the concave point and its two adjacent points before and after it. The distance between two concave points of the same adhesion is shorter than the distance from them to other concave points that satisfy the slope condition. Based on the above principles, the pseudo code of Concave Segmentation is shown in Algorithm 1.

Figure 10 shows the segmentation process for a complex adhesion cluster. After segmentation, the contour list of individual droplets can be directly obtained using the *findContours* and *contourArea* functions in *OpenCV*. Then, the deposition area of individual droplets in the contour list can be calculated and converted into volume. Further calculations can produce complex droplet measurement indicators such as *VMD* and *NMD*.



(**a**) (**b**) (**c**) (**d**)

Figure 10. Applicable scenarios in and effects of concave segmentation. (**a**) Complex adhesion patch. (**b**) Find contours. (**c**) Locate the concave point. (**d**) Concave point matching and segmentation (Colors are used to distinguish individuals).

Algorithm 1 Concave Segmentation algorithm

Input: Droplet adhesion patch image $I_{adherent}$
Output: Droplet individual patch image $I_{individual}$
Search for contours in droplet patch images
$L_{contours} \leftarrow$ findContours($I_{adherent}$)
Obtain corner point list using contour corner detection
$L_{corners} \leftarrow$ detectCorners ($L_{contours}$)
Locate the corner point and its two adjacent points before and after in the outline, and obtain 3 coordinate arrays.
L_{corner_locs}, L_{prev_locs}, $L_{next_locs} \leftarrow$ getLocations ($L_{corners}$)
Concave point pairing.
completed_pnts $\leftarrow$ [] # Record the paired droplets
$I_{individual} \leftarrow I_{adherent}$
for corner_loc_1, prev_loc_1, next_loc_1 $\in$ (L_{corner_locs}, L_{prev_locs}, L_{next_locs}) do
    if corner_loc_1 $\in$ completed_pnts then
        continue
    min_dist_corner_loc = null
    min_distance $\leftarrow\infty$
    for corner_loc_2, prev_loc_2, next_loc_2 $\in$ (L_{corner_locs}, L_{prev_locs}, L_{next_locs}) do
        if corner_loc_2 $\in$ completed_pnts then
            continue
        dintance $\leftarrow$ getDistance(corner_loc_1, corner_loc_2)
        if dintance < min_distance then
            if line_through_triangle(corner_loc_1, prev_loc_1, next_loc_1,
                corner_loc_2, prev_loc_2, next_loc_2) then
                min_dist_corner_loc $\leftarrow$ corner_loc_2
                min_distance $\leftarrow$ dintance
    if min_dist_corner_loc is not null then
        completed_pnts $\leftarrow$ completed_pnts $\cup$ corner_loc_1
        completed_pnts $\leftarrow$ completed_pnts $\cup$ min_dist_corner_loc
        line($I_{individual}$, corner_loc_1, corner_loc_2)
return $I_{individual}$

2.6. Droplet Segmentation Area Correction and Deposition Parameter Acquisition

Calculating *VMD* and *NMD* requires traversing individual droplets, so identifying individual droplets is the most basic prerequisite for calculating these two indicators. When the independence rate is regarded as the most important metric in the model, the accuracy of the droplet segmentation area may not be in the optimal state, resulting in statistical deviations in *VMD*, *NMD*, and global coverage. Therefore, it is necessary to correct the inaccurate segmentation area results.

The semantic segmentation model is equivalent to a function that implements mapping the image to the segmentation mask. It contains the rules for segmenting droplets with the model, which will be reflected in the segmentation results of the model. Finding this rule can correct the segmentation area of individual droplets, and improve the effectiveness of spray evaluation indicators such as *VMD*, *NMD*, and global coverage to a certain extent.

In the practical application of the method, the predicted area of the droplets is the only known quantity. If there is a mapping relationship between the predicted area A_{pred} of the droplet and the real area A_{true} of the droplet, deviation correction can be directly implemented through this relationship. According to experimental statistics, this relationship is not a fixed ratio, but changes with the size of the droplet. The smaller the size of the droplet, the greater the ratio $\gamma = \frac{A_{pred}}{A_{true}}$. When the droplet reaches a certain level, γ approaches a certain value. In addition, if the relationship between A_{pred} and A_{true} is directly fitted, the deviation in the area of small droplets will be ignored. The range of the area ratio γ's value is [0, 1], which is more accurate for small droplets when fitting the relationship and is more suitable as the dependent variable in the relationship function. If $\gamma = f_{\gamma}(A_{pred})$ is used to

represent the mapping function from the predicted area of the droplet to the area ratio γ, the deviation value can be expressed by Equation (7).

$$\delta = \frac{A_{pred}}{f_\gamma\left(A_{pred}\right)} - A_{pred} \tag{7}$$

The individual area of droplets corrected by Formula (6) is used to calculate droplet deposition parameters, including droplet density (D_t), droplet coverage (Dr), VMD, NMD, etc., which can be represented by Equations (8)–(11).

$$D_t = \frac{n}{S_{img}} \tag{8}$$

where n denotes the number of droplets in the sample, and S_{img} denotes the total area of the spray deposition image.

$$D_r = \frac{\sum_{i=1}^{n}\left(\frac{A_i}{f_\gamma(A_i)}\right)}{S_{img}} \tag{9}$$

where n denotes the number of droplets in the sample, A_i denotes the predicted area of the i-th droplet, f_γ denotes the correction formula from the predicted area to the real area, and S_{img} denotes the total area of the spray deposition image.

$$VMD = \sqrt[3]{\frac{6 * f_{vd}(L_{ads})}{\pi}} \tag{10}$$

where, L_{ads} denotes an array of area values sorted from small to large for all droplets in a single image sample, while f_{vd} denotes a function that includes logical judgment. It converts the input ordered area array into a volume sequence and calculates the volume of the droplet when the cumulative values of the elements in the volume sequence from small to large reach 50% of the total volume sequence.

$$NMD = \sqrt[3]{\frac{6 * f_{nd}(L_{ads})}{\pi}} \tag{11}$$

where L_{ads} denotes the area values of all droplets in a single image sample sorted from small to large, while f_{nd} denotes the volume calculated from the median element of the area array.

3. Results and Discussion

3.1. Results

3.1.1. Semantic Segmentation Result of DSRSS Model

In order to determine the optimal α and β in the $WIoU$ loss function, the model was trained on the DDID under different α and β values. The convergence curve of the model is shown in Figure 11. The performance indicators that the model can achieve when training for 100 epochs are shown in Figure 12. Each subgraph represents the optimal values that the model can achieve for each metric when training for 100 epochs under a pair of parameters α and β. The more balanced these optimal values are, the larger the area of the corresponding hexagonal shape enclosed by the points, indicating the better performance of the model. Figure 12 shows that the overall performance of the model reaches its optimum when α and β are 0.775 and 0.225, respectively. At this time, the IoU, recognition rate T_r, independence rate I_r, and correct rate C_r of the model are relatively high, and Precision and Recall are relatively balanced.

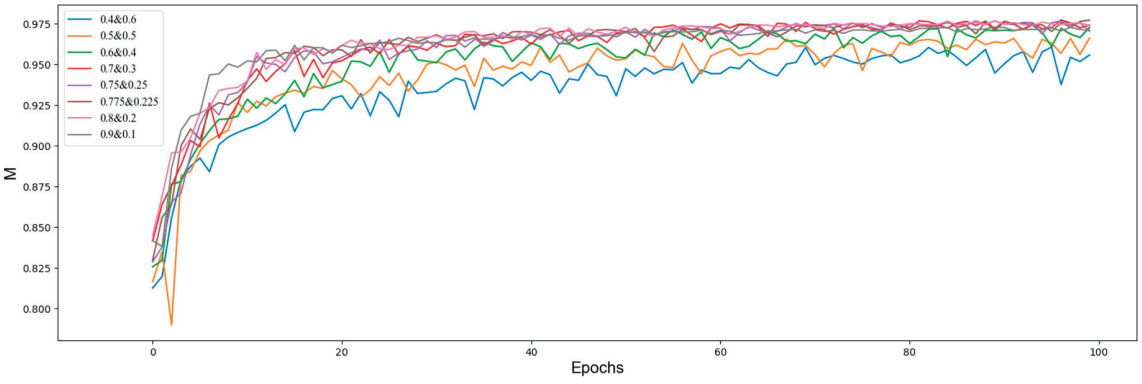


Figure 11. Convergence curve of the model under different weight values of the *WIoU* loss function (α and β, 100 epochs).

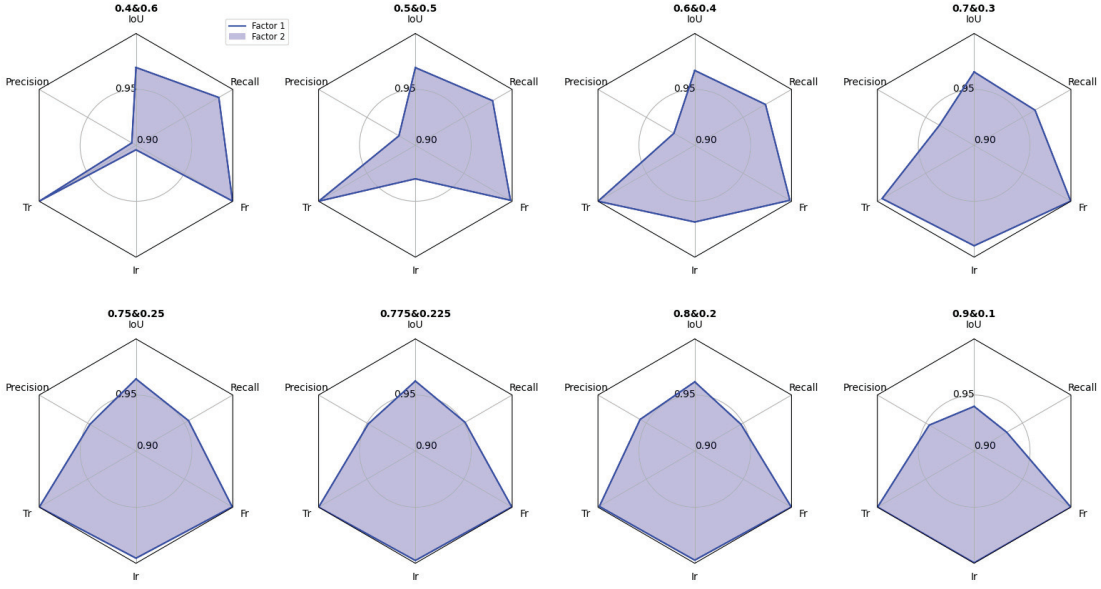


Figure 12. Optimal performance of the model under different weight values of the *WIoU* loss function (α and β, 100 epochs).

After determining the optimal α and β, the DSRSS model was trained more fully. The *IoU*, *Precision*, *Recall*, *Tr*, *Ir*, and *Cr* change curves of the model are shown in Figure 13. It shows that the convergence of *IoU*, Precision, and Recall is very smooth, while the convergence of *Tr*, *Ir*, and Cr is more turbulent, and the independence rate oscillates far more than other indicators. Therefore, the independence rate is used as the only indicator of the optimal weight value. In 1000 epochs of training, the independence rate of the model on the verification set reached 0.998, the recognition rate and correct rate both reached close to 1.0, and the *IoU* reached 0.973. Figure 14 shows the model's segmentation effect on the DDID verification sample. The comparison of the sample image with the colored segmentation mask shows that the segmentation of the edge of droplet is accurate, but the droplet independence needs to be improved.

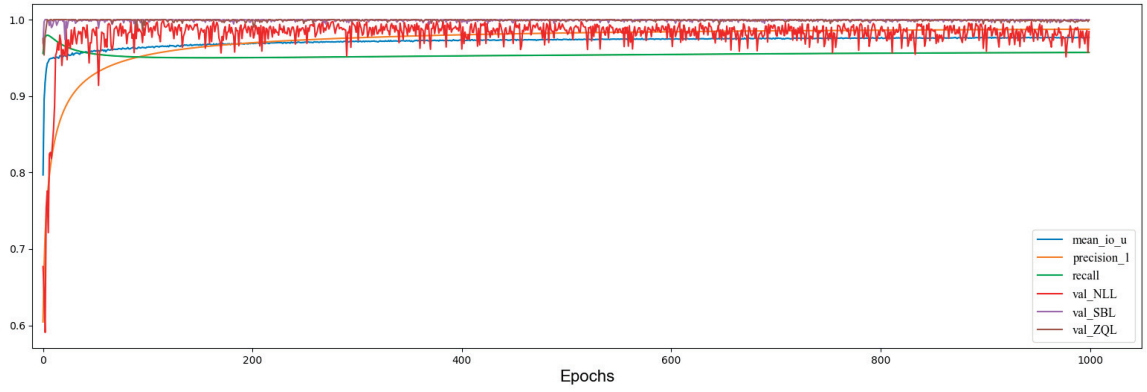
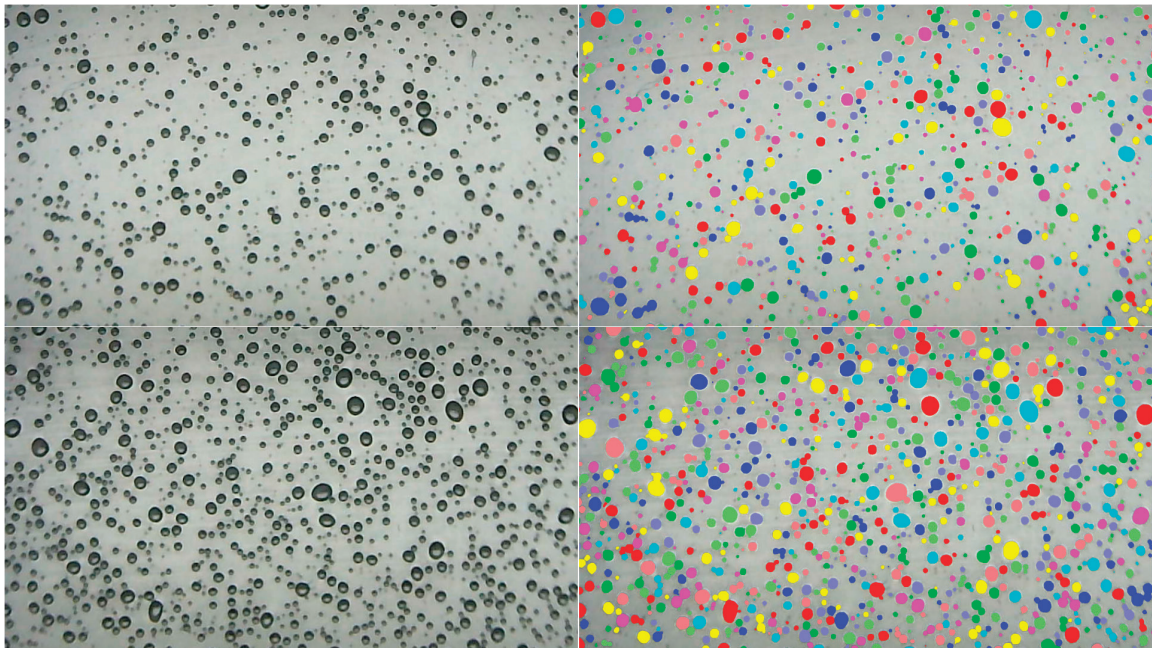


Figure 13. Full convergence of the model.

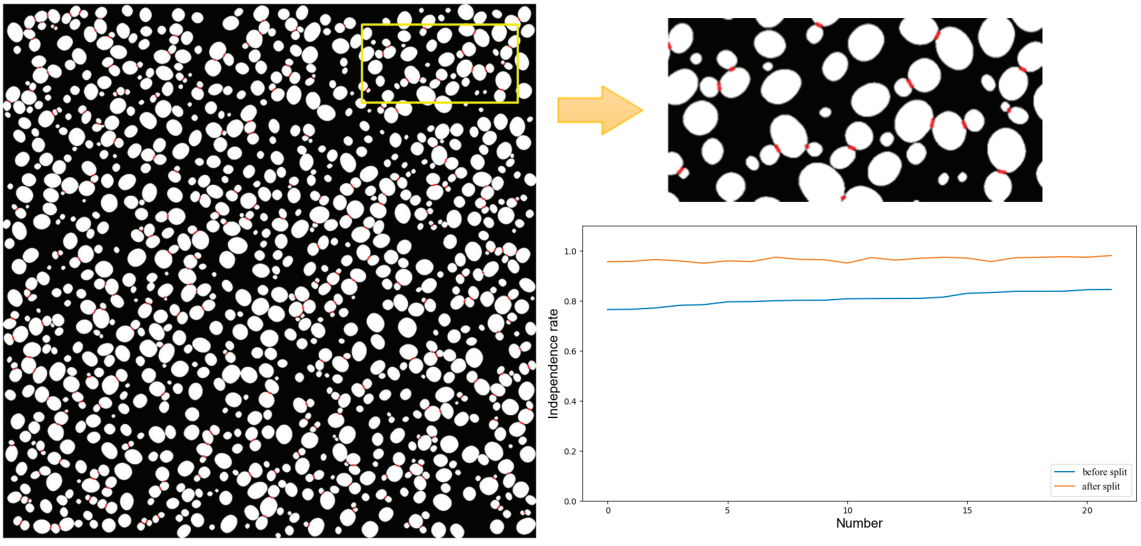


(**a**) Droplet deposition image (**b**) Model segmentation results

Figure 14. Segmentation results for fully trained model.

3.1.2. Adhesion Segmentation Result of MACS Algorithm

In order to verify the effect of the MACS algorithm, samples with poor quality, dense droplets, and severe adhesion were selected from the validation dataset set for testing. An example of the selected image sample is shown in Figure 15a. The results show that the auxiliary segmentation algorithm can further improve the independence rate of droplet segmentation. The average independence rate of 22 samples increased by 15.7%, as shown in Figure 15b.



(**a**) A severely adhesive sample (**b**) Comparison of independence rates before and after MACS segmentation

Figure 15. Testing of MACS algorithm.

3.1.3. Result of Droplet Segmentation Area Correction

The droplet areas predicted by the model for the validation dataset and the ratio of these predicted areas to the corresponding labeled areas were counted. After removing incorrectly segmented droplets and adherent droplets, a total of 98,795 pairs of original data were obtained. The distribution of the predicted droplets was divided into 256 intervals, and all data in each interval were averaged, resulting in a total of 256 pairs of mean sample data. Using 256 pairs of mean data, the conversion formula Equation (12) was fitted. The R2 value of Equation (12) and 256 mean sample points reached 99.07. The distribution of original sample pairs, the distribution of mean sample pairs, and the graph of Equation (12) are shown in Figure 16.

$$\gamma = \frac{1}{\left(1 + 24.0717268^{-A_{pred}^{0.133783199}}\right)^{73.9259088}} - 0.0196201541 \quad (12)$$

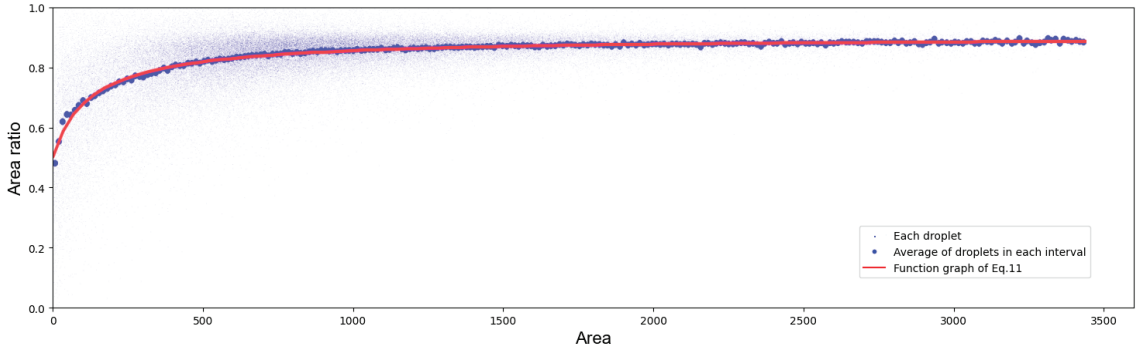


Figure 16. γ Distribution and Function Fitting.

In order to verify the effect of Equation (12) in global indicator statistics, the validated samples were selected to calculate three coverage rates. They are the ground truth coverage,

the coverage directly calculated from the model output, and the model's output coverage corrected by Equation (12). Each group of data were arranged from large to small according to the ground truth coverage, and the line chart drawn is shown in Figure 17. The correction function reduces the coverage error from 4.67% to 1.12%, showing that the correction function can improve the global statistical indicators.

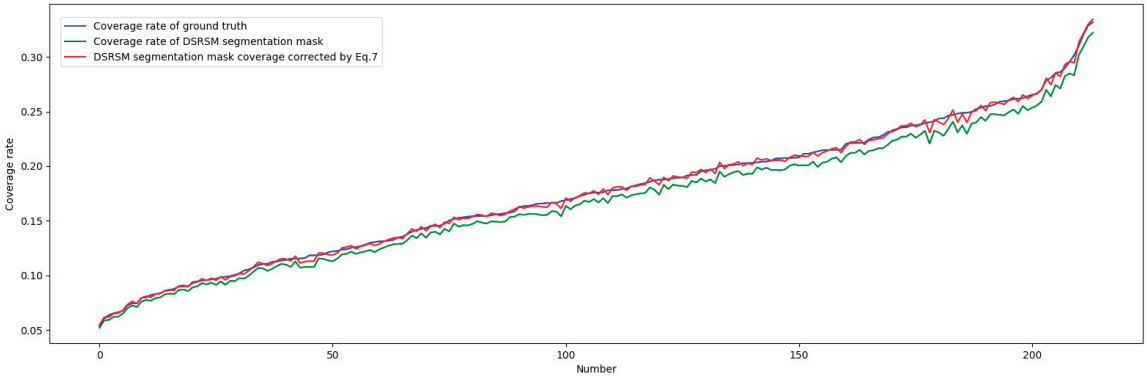
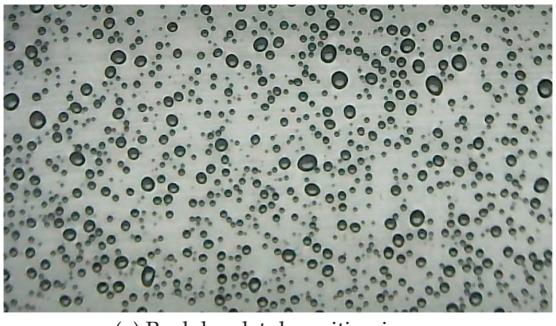


Figure 17. The effect of correction function on coverage rate.

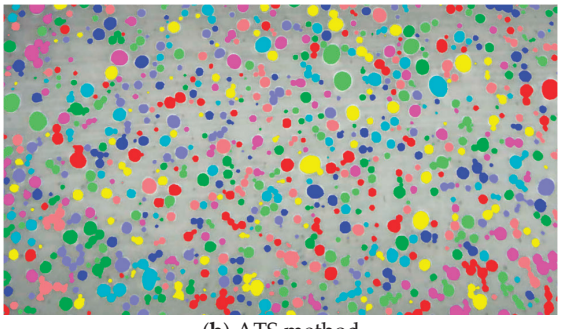
3.1.4. Result of Real Sample Processing

Finally, we input the real droplet deposition images captured by the device into the model, and the segmentation mask obtained after inference is shown in Figure 18. Perceptually, the segmentation performance of the model is significantly better than that of the adaptive threshold segmentation (ATS) method. The vast majority of droplets are effectively segmented, including those that are adhered during threshold segmentation, and the obtained edges are smoother than threshold segmentation, filling almost all droplet bright spots. The above advantages are also reflected in the spray parameters; Table 1 shows the spray parameters calculated based on the segmentation masks output by the two methods, including the coverage, droplet number, droplet density, and VMD. Comparing the segmentation boundaries of droplets in the two segmentation masks in Figure 18, the coverage obtained by our method is closer to the real situation, and due to the higher droplet independency rate, our method outputs a higher droplet density and a smaller value VMD value.

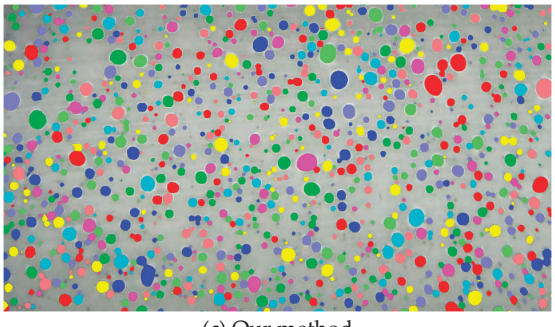


(**a**) Real droplet deposition image

Figure 18. *Cont.*



(**b**) ATS method



(**c**) Our method

Figure 18. Comparison of performance between DSRSS and ATS methods and the real droplet deposition image.

Table 1. Comparison of spray parameters between DSRSS and ATS methods.

Segmentation Method	Coverage Rate	Number of Droplets	Droplet Density (n/mm^2)	VMD (μm)
ATS	0.220	835	202	98.1
Our method	0.176	1044	252	76.8

3.1.5. Ablation Experiment of the Structure of the DSRSS Model

To verify the necessity of each component in the model, ablation experiments were performed. Figure 19 shows the performance of the model under four conditions: unchanged, downsampling reduced by one level, halved width, and cross layer connection removed. CLearly, the default structure model has the better performance.

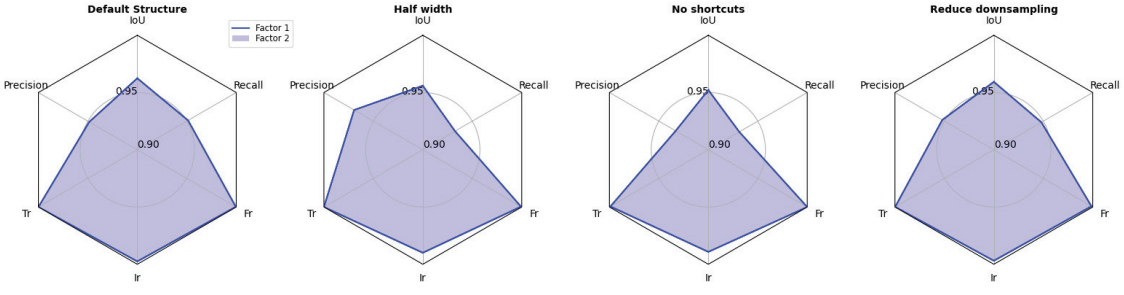


Figure 19. Performance of the model with changed structure.

3.2. Discussion

3.2.1. Rationality of Semantic Segmentation

From the perspective of the form of the task, target detection, instance segmentation, and semantic segmentation based on deep learning can all be applied to the segmentation and statistics of droplet images, replacing some processes in spray measurement work, as shown in Figure 20. The colored parts in the figure represent the work that the corresponding method can accomplish.

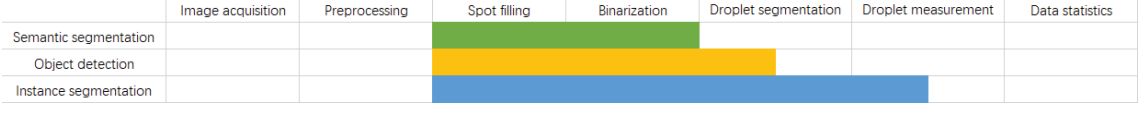
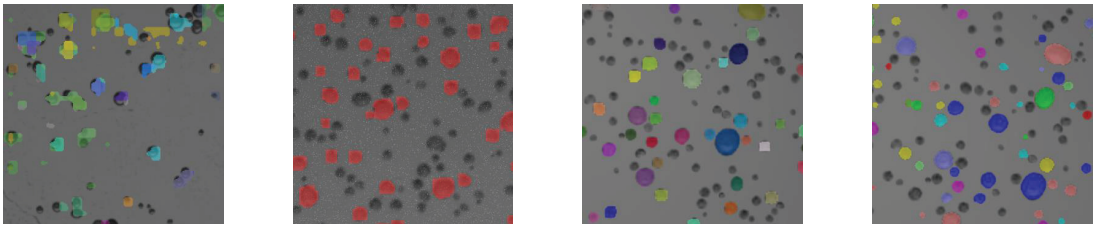


Figure 20. The role of three types of depth learning methods in spray measurement.

Among the above three types of methods, the instance segmentation algorithm can directly output the mask of individual droplets and realize the direct calculation of the deposition area, which seems to be the most reasonable choice. However, instance segmentation of both small objects and dense objects is extremely challenging. Perhaps because the droplets of high-quality spray are too small and too dense, in our tests, several instance segmentation models failed to show an acceptable segmentation accuracy, as shown in Figure 21.



(**a**) (**b**) (**c**) (**d**)

Figure 21. Performance of four instance segmentation methods in droplet segmentation. (**a**) YOLACT [22]. (**b**) YOLOv8 [23]. (**c**) FastSAM [24]. (**d**) SAM [25].

Since the segmentation task only needs to segment a single type of object, all other areas in the image except the droplets are considered as background. First, we performed semantic segmentation on the image, obtained the segmentation mask, and then segmented the individual droplets through an additional process, which is equivalent to using the instance segmentation task on a single type of object. Moreover, DSRSS already has a certain ability to focus on individual droplets. If the image is enlarged so that the size of the droplets is far beyond the reasonable range, and then input into the model, the result shown in Figure 22 is obtained. This shows that the display model is able to "consciously" reconstruct the shape and contour of the droplet from the blurred image based on the prior knowledge it learned.

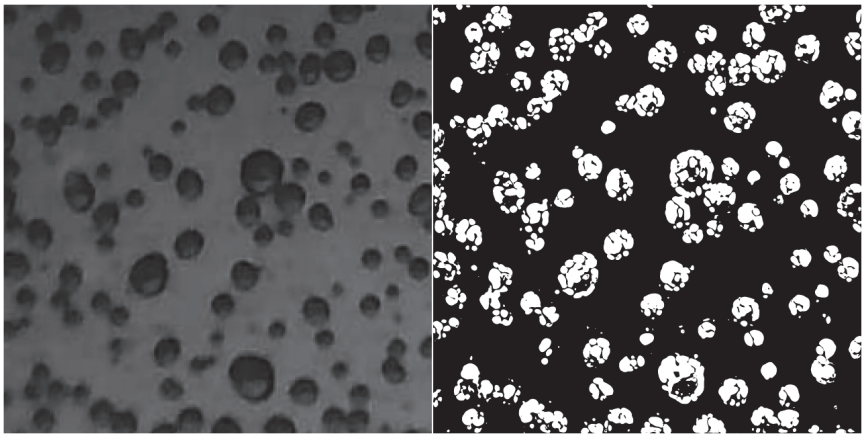


Figure 22. The segmentation effect of the model when the droplet size is too large.

3.2.2. Computing Power Requirements for the DSRSS Model

The number of parameters in the DSRSS model is less than 1 M. Since it is a fully convolutional network, it can accept inputs of different sizes. Since the encoder part of the model contains three times of downsampling, the length and width of the input image are required to be an integral multiple of eight. The computational complexity and RAM requirements of the model are positively correlated with the input size. Tests were

conducted on several computer hardware and the RAM capacity consumed by inferring single images of different sizes was recorded, the amount of computation, and the inference speed on different computing hardware. The results are shown in Table 2. Since the spray quality evaluation work does not require absolute real-time performance, its inference speed is tolerable on typical devices.

Table 2. Performance of models on different hardware platforms under different input sizes.

Input Size (Pixels)	RAM Consumption (GiB)	Calculation Quantity (GFLOPS)	Inference Speed (FPS)			
			Nvidia Tesla P40	Nvidia GTX 1660Ti Max-Q	Intel Core i7-9750H	Intel Xeon E3-1245V5
64 × 64	0.4	0.99	51.3	46.1	28.6	24.3
128 × 128	0.5	3.95	49.8	40.4	12.5	11.5
256 × 256	0.7	15.8	42.3	37.2	4.16	3.86
384 × 384	1.0	35.6	34.9	26.0	2.03	1.88
512 × 512	1.3	63.2	24.8	16.3	1.12	1.08

4. Conclusions and Future Work

The paper provides a lightweight end-to-end DSRSS model and a MACS algorithm to improve the accuracy of semantic segmentation in droplet deposition images and achieve precise and efficient multi-parameter measurement of droplet deposition with low-quality droplet deposition images for a scalable low-cost spraying measurement system for use in field scenarios. Based on DSRSS, MACS, and droplet area correction, a complete image processing procedure for droplet deposition images in a low-cost spray measurement system has been established. The main conclusions are as follows.

(1) When the two weight parameters α and β in $WIoU$ are 0.775 and 0.225, respectively, the rate of independent droplet segmentation when using the DSRSS on the DDID reaches 0.998 and the IoU reaches 0.973.
(2) The MACS algorithm can reduce the adhesion rate of droplet deposition images with a coverage exceeding 30% by 15.7%, and the area correction function can reduce the area segmentation error in the DSRSS by 3.54%.
(3) The quantity of parameters in the DSRSS is less than 1 M, and it can reduce RAM requirements through image input blocking, thus being able run on embedded platforms.
(4) Based on the exact test metrics obtained from the test dataset and the visualization of segmentation in real droplet deposition images, the comprehensive performance of this method is much better than that of the adaptive threshold segmentation method, as its droplet boundary segmentation is more accurate and the accuracy of droplet area estimation is higher.

This work can be used as a solution for high-speed in situ spray measurement, and provides a reference for the segmentation of low-quality dense object images. The complexity of the in situ measurement environment, device aging, and wear, especially the aging of the light source and the wear of the droplet collector, as well as the residue on the droplet collector after long-term use, will all affect the actual effectiveness of the method. As the next step, we plan to further improve the robustness of the method and make it suitable for more complex application scenarios.

Author Contributions: J.L.: Writing—original draft, Methodology, Data curation, Validation, Software. S.Y.: Investigation, Validation. X.L.: Writing—Review and Editing, Conceptualization, Methodology, Funding acquisition. G.L.: Data curation, Validation. Z.X.: Methodology, Investigation, Validation. J.Y.: Investigation, Resources, Project administration, Supervision. All authors have read and agreed to the published version of the manuscript.

Funding: This research was supported by National Natural Science Foundation of China (52075308) and National Key R&D Program of China (2022YFD2300101) and Shandong Provincial Cotton Industry Technology System Innovation Team Machinery Post Expert Project (SDAIT-03-09).

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: Anyone can access the data by sending an email to jyuan@sdau.edu.cn.

Conflicts of Interest: The authors declare no conflicts of interest.

References

1. Makwana, A.; Mohnot, P. Effect of spraying using sprayer robot for cotton crop: Sprayer robot for cotton crop. *J. AgriSearch* **2022**, *9*, 255–259. [CrossRef]
2. Guo, N.; Liu, S.; Xu, H.; Tian, S.; Li, T. Improvement on image detection algorithm of droplets deposition characteristics. *Trans. Chin. Soc. Agric. Eng.* **2018**, *34*, 176–182. [CrossRef]
3. Sarghini, F.; Visacki, V.; Sedlar, A.; Crimaldi, M.; Cristiano, V.; de Vivo, A. First measurements of spray deposition obtained from UAV spray application technique. In Proceedings of the 2019 IEEE International Workshop on Metrology for Agriculture and Forestry (MetroAgriFor), Portici, Italy, 24–26 October 2019; pp. 58–61. [CrossRef]
4. Wang, P.; Yu, W.; Ou, M.; Gong, C.; Jia, W. Monitoring of the pesticide droplet deposition with a novel capacitance sensor. *Sensors* **2019**, *19*, 537. [CrossRef] [PubMed]
5. Wen, Y.; Zhang, R.; Chen, L.; Huang, Y.; Yi, T.; Xu, G.; Hewitt, A.J. A new spray deposition pattern measurement system based on spectral analysis of a fluorescent tracer. *Comput. Electron. Agric.* **2019**, *160*, 14–22. [CrossRef]
6. Yang, W.; Li, X.; Li, M.; Hao, Z. Droplet deposition characteristics detection method based on deep learning. *Comput. Electron. Agric.* **2022**, *198*, 107038. [CrossRef]
7. Zhu, H.; Salyani, M.; Fox, R.D. A portable scanning system for evaluation of spray deposit distribution. *Comput. Electron. Agric.* **2011**, *76*, 38–43. [CrossRef]
8. Wang, L.; Song, W.; Lan, Y.; Wang, H.; Yue, X.; Yin, X.; Tang, Y. A Smart Droplet Detection Approach With Vision Sensing Technique for Agricultural Aviation Application. *IEEE Sens. J.* **2021**, *21*, 17508–17516. [CrossRef]
9. Vacalebre, M.; Frison, R.; Corsaro, C.; Neri, F.; Santoro, A.; Conoci, S.; Anastasi, E.; Curatolo, M.C.; Fazio, E. Current State of the Art and Next Generation of Materials for a Customized Intraocular Lens according to a Patient-Specific Eye Power. *Polymers* **2023**, *15*, 1590. [CrossRef]
10. Gennes, P.-G.; Brochard-Wyart, F.; Quéré, D. *Capillarity and Wetting Phenomena: Drops, Bubbles, Pearls, Waves*; Springer: Berlin/Heidelberg, Germany, 2004.
11. Kupyn, O.; Budzan, V.; Mykhailych, M.; Mishkin, D.; Matas, J. 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, 18–23 June 2018. [CrossRef]
12. Kupyn, O.; Martyniuk, T.; Wu, J.; Wang, Z. Deblurgan-v2: Deblurring (orders-of-magnitude) faster and better. In Proceedings of the IEEE/CVF International Conference on Computer Vision, Seoul, Republic of Korea, 27 October–2 November 2019.
13. Quan, Y.; Wu, Z.; Ji, H. Gaussian kernel mixture network for single image defocus deblurring. *Adv. Neural Inf. Process. Syst.* **2021**, *34*, 20812–20824. [CrossRef]
14. Ledig, C.; Theis, L.; Huszár, F.; Caballero, J.; Cunningham, A.; Acosta, A.; Aitken, A.; Tejani, A.; Totz, J.; Wang, Z.; et al. Photo-realistic single image super-resolution using a generative adversarial network. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Honolulu, HI, USA, 21–26 July 2017. [CrossRef]
15. Wang, L.; Li, D.; Zhu, Y.; Tian, L.; Shan, Y. Dual super-resolution learning for semantic segmentation. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, Seattle, WA, USA, 13–19 June 2020.
16. Wang, X.; Yu, K.; Wu, S.; Gu, J.; Liu, Y.; Dong, C.; Change Loy, C. Esrgan: Enhanced super-resolution generative adversarial networks. In Proceedings of the European Conference on Computer Vision (ECCV) Workshops, Munich, Germany, 8–14 September 2018. [CrossRef]
17. Ronneberger, O.; Fischer, P.; Brox, T. U-net: Convolutional networks for biomedical image segmentation. In Proceedings of the Medical Image Computing and Computer-Assisted Intervention–MICCAI 2015: 18th International Conference, Munich, Germany, 5–9 October 2015. [CrossRef]
18. Badrinarayanan, V.; Kendall, A.; Cipolla, R. Segnet: A deep convolutional encoder-decoder architecture for image segmentation. *IEEE Trans. Pattern Anal. Mach. Intell.* **2017**, *39*, 2481–2495. [CrossRef] [PubMed]
19. Chen, L.-C.; Zhu, Y.; Papandreou, G.; Schroff, F.; Adam, H. Encoder-decoder with atrous separable convolution for semantic image segmentation. In Proceedings of the European Conference on Computer Vision (ECCV), Munich, Germany, 8–14 September 2018. [CrossRef]
20. Long, J.; Shelhamer, E.; Darrell, T. Fully convolutional networks for semantic segmentation. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Boston, MA, USA, 7–12 June 2015. [CrossRef]
21. Ramachandran, P.; Zoph, B.; Le, Q.V. Searching for Activation Functions. *arXiv* **2017**, arXiv:1710.05941. [CrossRef]
22. Bolya, D.; Zhou, C.; Xiao, F.; Lee, Y.J. Yolact: Real-time instance segmentation. In Proceedings of the IEEE/CVF International Conference on Computer Vision, Seoul, Republic of Korea, 27 October–2 November 2019.

23. Terven, J.; Cordova-Esparza, D. A Comprehensive Review of YOLO: From YOLOv1 to YOLOv8 and Beyond. *arXiv* **2023**, arXiv:2304.00501. [CrossRef]
24. Zhao, X.; Ding, W.; An, Y.; Du, Y.; Yu, T.; Li, M.; Wang, J. Fast Segment Anything. *arXiv* **2023**, arXiv:2306.12156. [CrossRef]
25. Kirillov, A.; Mintun, E.; Ravi, N.; Mao, H.; Rolland, C.; Gustafson, L.; Xiao, T.; Whitehead, S.; Berg, A.C.; Lo, W.-Y.; et al. Segment Anything. *arXiv* **2023**, arXiv:2304.02643.

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.

agriculture

MDPI

Article

Design and Validation of a Variable-Rate Control Metering Mechanism and Smart Monitoring System for a High-Precision Sugarcane Transplanter

Abdallah E. Elwakeel [1,*], Yasser S. A. Mazrou [2], Ahmed S. Eissa [3], Abdelaziz M. Okasha [4], Adel H. Elmetwalli [5], Abeer H. Makhlouf [6], Khaled A. Metwally [7], Wael A. Mahmoud [8] and Salah Elsayed [9,10]

1 Agricultural Engineering Department, Faculty of Agriculture and Natural Resources, Aswan University, Aswan 81528, Egypt
2 Applied College, Muhayil Asir, King Khalid University, Abha 62587, Saudi Arabia; ymazrou@kku.edu.sa
3 Agricultural Products Process Engineering Department, Faculty of Agricultural Engineering, Al-Azhar University, Cairo 11751, Egypt; ahmedeissa2205.el@azhar.edu.eg
4 Department of Agricultural Engineering, Faculty of Agriculture, Kafrelsheikh University, Kafr El-Sheikh 33516, Egypt; abdelaziz.okasha@agr.kfs.edu.eg
5 Agricultural Engineering Department, Faculty of Agriculture, Tanta University, Tanta 31527, Egypt; adel.elmetwali@agr.tanta.edu.eg
6 Department of Agricultural Botany, Faculty of Agriculture, Minufiya University, Shibin El-Kom 32511, Egypt; abeer.makhlof@agr.menofia.edu.eg
7 Soil and Water Sciences Department, Faculty of Technology and Development, Zagazig University, Zagazig 44511, Egypt; khametwally@zu.edu.eg
8 Department of Agricultural Machinery and Power Engineering, Faculty of Agricultural Engineering, Al-Azhar University, Assiut 71524, Egypt; waelshaban.50@azhar.edu.eg
9 Agricultural Engineering, Evaluation of Natural Resources Department, Environmental Studies and Research Institute, University of Sadat City, Sadat City 32897, Egypt; salah.emam@esri.usc.edu.eg
10 New Era and Development in Civil Engineering Research Group, Scientific Research Center, Al-Ayen University, Nasiriyah 64001, Thi-Qar, Iraq
* Correspondence: abdallah_elshawadfy@agr.aswu.edu.eg

Citation: Elwakeel, A.E.; Mazrou, Y.S.A.; Eissa, A.S.; Okasha, A.M.; Elmetwalli, A.H.; Makhlouf, A.H.; Metwally, K.A.; Mahmoud, W.A.; Elsayed, S. Design and Validation of a Variable-Rate Control Metering Mechanism and Smart Monitoring System for a High-Precision Sugarcane Transplanter. *Agriculture* **2023**, *13*, 2218. https://doi.org/10.3390/agriculture13122218

Academic Editors: Jin Yuan and Zichen Huang

Received: 19 October 2023
Revised: 15 November 2023
Accepted: 25 November 2023
Published: 30 November 2023

Copyright: © 2023 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (https://creativecommons.org/licenses/by/4.0/).

Abstract: The current study aimed to design and test the accuracy of a variable-rate control metering mechanism (VRCMM) and a remote smart monitoring system (RSMS) for a precision sugarcane transplanter based on IoT technology. The VRCMM is used to operate the seedling metering device at different speeds using a stepper motor based on the travel speed, and the RSMS was employed to evaluate of the three basic parameters of seedling amount, optimum rate, and missed rate. Two types of sensors were used for detecting the sugarcane seedling (SS) and travel speed, including one ultrasonic sensor and one infrared RPM sensor. The study was performed at five travel speeds and four transplant spacings. The findings of laboratory tests showed that the mean record of the relative error between the desired stepper motor speed of the VRCMM and the real value was 3.39%, and it increased with increasing the travel speed. as Additionally, the speed regulation performance was in agreement with the transplanting index. The change in RSMS accuracy is obvious when the travel speed is high and the transplant spacing is small. The RSMS accuracy drops sharply, revealing a leaping change. In conclusion, the smart and intelligent designed sugarcane transplanter would be very useful in sugarcane production.

Keywords: planter; seeder; metering device; ultrasonic sensor; Arduino; precision agriculture

1. Introduction

Worldwide, sugar is regarded as one of the most important strategic commodities, and in Egypt, it is ranked the second most important strategic commodity after wheat. Thus, it was essential to give our whole attention to sugarcane for enhancing crop yield and to

decrease the gap between consumption and crop production. In Egypt, sugarcane is grown on over 126,000 ha of agricultural land having an average production of 114.3 ton/ha, and total annual production may reach 16 million tons [1–5].

Planting technique represents one of the most crucial practices in the production of sugarcane. It results in a quickly and consistently emerging plant stand at the desired plant density. However, it takes a lot of time and work [6–8]. Mechanized field planting has become more popular in the sugarcane industry as a solution to labor shortages and expensive production [9]. Rípoli et al. [10] compared five different sugarcane planters' cost effectiveness to semi-mechanized equipment and stated that the semi-mechanized equipment was much more expensive. In addition, automatic sugarcane planting machine costs less to operate. It greatly saves labor in the working process, and operates extended hours during both day and night shifts [11]. At this time, the three primary types of sugarcane planter are whole-stalk planters, real-time planters, and billet planters [6]. In addition to other industrialized countries, the United States, Japan, and Italy, among others, have begun to investigate sugarcane transplanting and have created automatic transplanter machinery suitable for diverse geographical environments and crop varieties [12–15].

Precision agriculture (PA) is a methodical strategy for restructuring the production agricultural system in order to achieve a low-input, highly effective, and long-lasting system [16,17]. Developing a decision-support system for farm management with the goal of maximizing input return while preserving resources is PA's primary goal [18]. With the benefits of conserving seeds, lowering labor intensity, and optimizing operational efficiency, precision seeding and planting is going to be an essential direction for advancement in agriculture in the coming years [19,20]. PA is also a very important aspect for transplanting sugarcane at the optimum time, with the optimum moisture content, at the same intervals. Total grain seeds must be uniformly distributed throughout the field in addition to being planted or drilled at the ideal planting or seeding rate per area in order to produce the desired crop rows. Additionally, maintaining the ideal plant-to-plant spacing boosts crop yield [21–26]. Plant spacing and crop yield have a well-established relationship [27–29]. For numerous crops, row plant spacing is influenced just as significantly by the crop's specific biological spacing demands as it is by the structural design of the agricultural machines utilized [30]. In addition, a key point of the precise planting of a precision transplanting or seeding machine in a typical square grid pattern needs equipment calibration tests before planting in both directions of the crop row [31]. With the advancement of precision planting in smart agriculture, variable rate application (VRA), which aims to increase savings of fertilizers, seeds, and herbicides in keeping with site specific management particular administration basics within a field, has grown in popularity [32].

By modifying seeding rates based on field conditions, this crucial PA component has substantially raised agricultural production and preserved seeds while also raising yield and net farm returns [33,34].

The pace of mechanization in sugarcane planting is increasing daily. Over several decades, many researchers worldwide have created sugarcane billet measuring devices. A lot of research has been conducted regarding the uniformity of sugarcane billet planting machines. Razavi et al. [35] investigated how a sugarcane billet planter's (SBP) forward speed and the chain conveyor angle affected planting uniformity. Taghinezhad et al. [36] studied the impact of rotation angle and forward speed of SBPs on precision indexes and discharging. Saengprachatanarug et al. [37] improved the discharge mechanism of the SBP by adjusting the slope of the billet's hopper. Saengprachatanarug et al. [38] modified a metering mechanism for SBP by developing the arrangement of the cleat's conveyor. Moslem et al. [39] developed a metering mechanism capable of planting sugarcane billets at 50 cm spacing.

The foundation of the automation and intelligence of seeding and transplanting, as well as the trustworthy assurance of signal feedback and equipment research along with the development of the follow-up seeding links, has enhanced the efficiency of seeding and transplanting. Photoelectric monitoring, high-speed photography, and piezoelectric

sensing are just examples of monitoring techniques that many researchers have utilized to monitor seeding quality [26,40–43]. The most common criteria for evaluating seeding quality include seeding quantity, missed rate, and optimum rate. It is usually necessary to set up a monitoring device of seedlings in order to obtain these parameters. Furthermore, monitoring sensors are an extremely essential part of the monitoring device.

Currently, companies have created and designed different monitoring technologies such as laser diodes, optical sensors, photoresistors, and infrared sensors, as well as optical fiber sensors being examples of these technologies [21,44–46]. In addition, many researchers have employed image processing, piezoelectric sensors, and capacitive sensors to track the flow data of seeds moving via the seed tube [22,47,48]. Yet, there are no automated seedling sensors available for sugarcane or vegetable transplanters. To the best of the authors' knowledge, there are not yet many SS meters that can transplant varying amounts of sugarcane. In this regard, too much research is required to enhance sugarcane transplanting efficiency. Based on the abovementioned discussion, the current study provides technical and practical references for redesigning sugarcane precision transplanting–metering machines and also offers crucial machine innovation for the progress of precision agriculture.

The objectives of the study were to:

1. Design and assess the performance of a variable-rate control metering mechanism (VRCMM) for sugarcane transplanting based on IoT technology.
2. Design and test a remote smart monitoring system (RSMS) for a precise sugarcane transplanter based on an ultrasonic sensor, IoT technology, Android, and wireless communication.

2. Materials and Methods

2.1. SSs

A twenty-month-old C-9 sugarcane variety was used for this investigation. Applying a sugarcane bud chipping machine generated by Elwakeel et al. [5], sugarcane buds were cut off 1.5 m from the roots of healthy sugarcane plants that were free of insects and diseases. In a plastic greenhouse, the extracted sugarcane buds were planted in a plastic tray filled with farmyard manure (FYM), which is prepared basically using cow dung, cow urine, rice straw, and other crop residues. The FYM was mixed with soil at a ratio of 1:1 and grown. After 70 days of planting, the seedlings were at their ideal height of 35 cm for transplanting.

2.2. Overall Structure and Principles of Working

The overall structure of a variable-rate control metering mechanism (VRCMM) for the sugarcane transplanter is shown in Figure 1, and was composed of a seedling cylinder, a stepper motor, its driver, a pinion, a sprocket, a speed sensor, an Arduino mega board (AMB), and a seedling counter sensor (ultrasonic sensor). As the seedling cylinder rotated clockwise, the SSs in the rotating seedling cylinders would fill the seedling-guiding slot (SGS), and then be discharged on the earlier open ridge. To track the number of discharged seedlings, a remote smart monitoring system (RSMS) was installed at the outlet of the SGS.

As seen in Figure 1, the sugarcane transplanter is fundamentally composed of a two-wheel walking tractor and a transplanting unit. The Bluetooth unit and control unit are fixed on the upper point of the two-wheel walking tractor to receive and transmit signals for precise management, operation control, and remote monitoring. The transplanting unit is attached to the tractor via a four-point attaching and primarily consists of a furrow opener, a rotary seedling cylinder, two soil-compaction wheels, an RSMS, and a VRCMM.

2.3. Design of Electronic Circuits for VRCMM and RSMS

The VRCMM and the RSMS consist of six main parts: an infrared light speed sensor (model: LM 393 IC, China), an ultrasonic sensor (model: HC-SR04, Hammond Manufacturing L., Canada), an Arduino Mega (model: 2500 R3, China), a Bluetooth unit (model: HC-05, Yantel Corporation, China), stepper motors (model: Nema 23, CUI Devices, Tualatin, USA), and a stepper motor driver (model: TB 6600, Sorotec GmbH, Germany). In addition, some

axillary components, such as a laptop (model: TOSHIBA Intel (R) Core (TM) i3-3120M CPU @ 2.50 GHz, Tokyo) or smart phone (OPPO A15, OPPO, Guangdong, China), bread boards, wires, a USB cable, a PV panel (60 W, Eagle Network Supply Private Limited, India), a battery charger (12 V, Victron Energy, The Netherlands), a converter (12 V–5 V) {Optional}, and battery (70 A·h), were also used. These components are depicted and described in Figures 2–4. Figure 2 illustrates signal flowrate, electric energy connections, and mechanical connections between the different electronic parts. A schematic diagram of the VRCMM and the RSMS with successful electrical connections is shown in Figure 3. Figure 4 demonstrates the technical specifications of the main electronic components of the VRCMM and the RSMS.

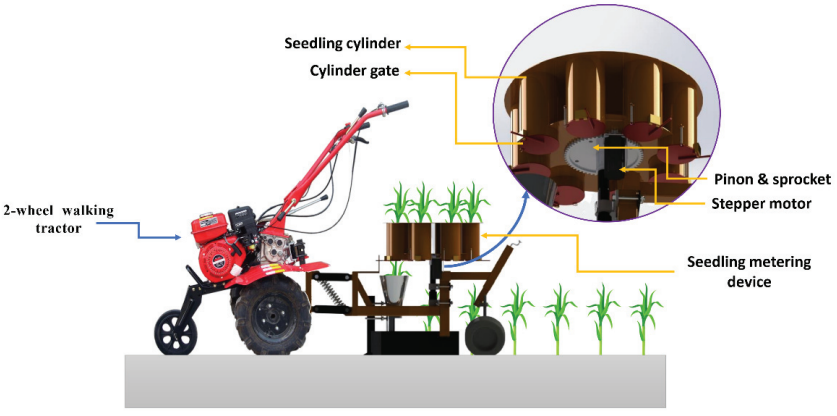


Figure 1. The overall structure of the VRCMM.

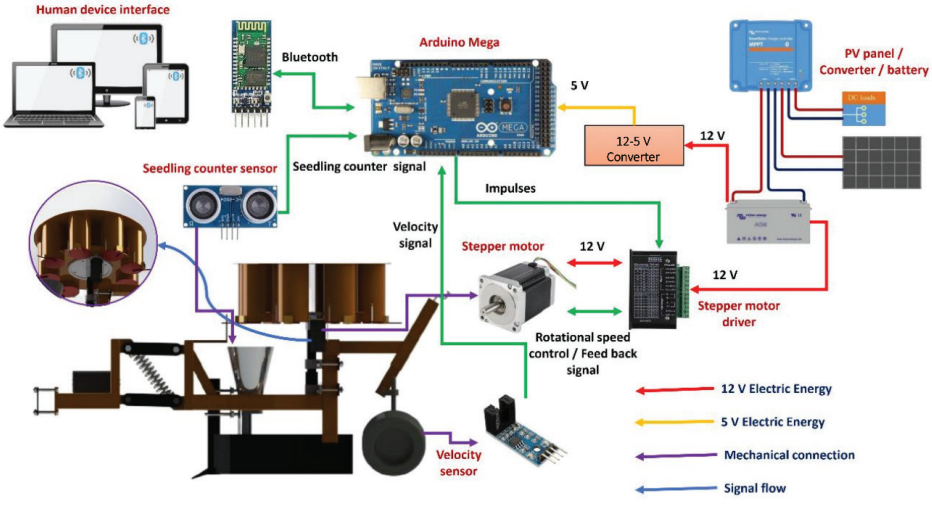


Figure 2. The operation map of the VRCMM and the RSMS.

2.3.1. Speed Sensor

The speed sensor (mode: LM393 IC) was employed to determine the ground wheel rotational speed. The speed sensor comprises of a phototransistor and an infrared LED which are fixed on a small, printed circuit board. When the LED is switched on, it releases a beam of infrared rays (IR). The phototransistor will detect any interruption of the IR beam caused by the rotating ground wheel, and produce an electrical pulse. The number of pulses per second is closely linked to the RPM of the ground wheel. In addition, using a microcontroller, it may be utilized to determine the number of pulses per second and then

convert that count to RPM, and finally to forward speed (linear speed). The speed sensor can work in a voltage range of 3.3 to 5 V. The output signal is high, and it is low when the sensor detects pulses. Figure 5 illustrates the main parts and connecting pins of the speed sensor used in the VRCMM.

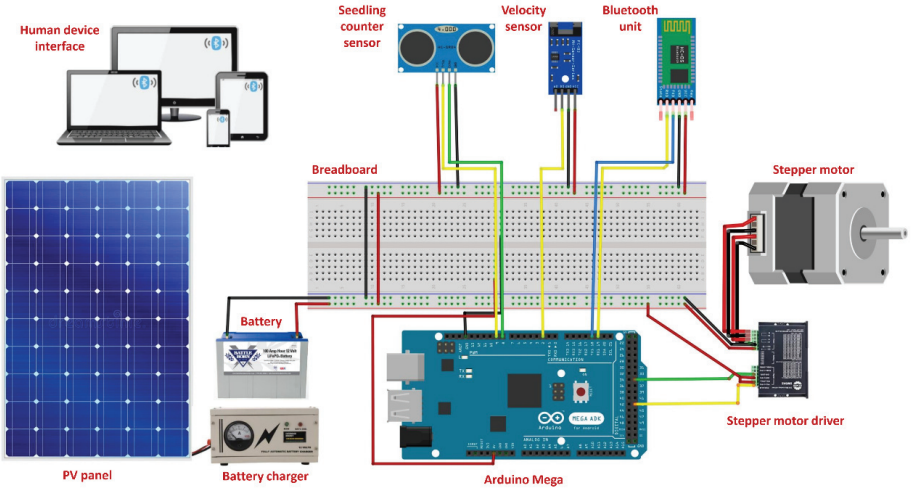


Figure 3. Schematic diagram of the VRCMM and the RSMS with successful electrical connections.

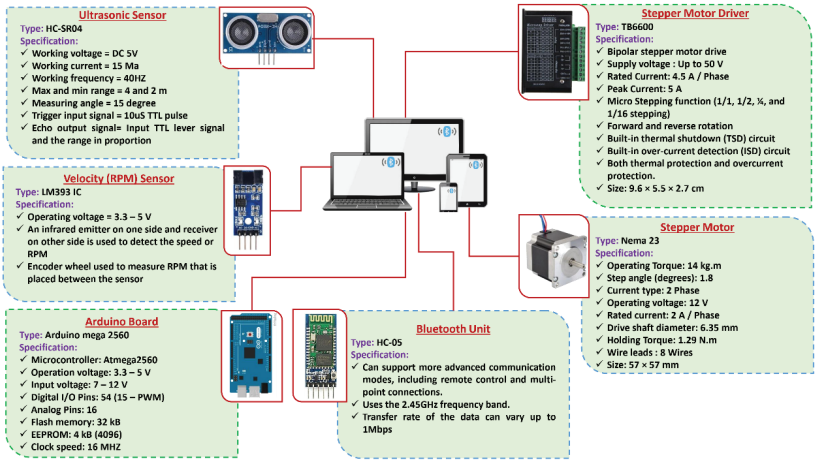


Figure 4. Specification of the main electronic components of the VRCMM and the RSMS.

2.3.2. Ultrasonic Sensor (Seedling Counter Sensor)

The ultrasonic sensor (model: HC-SR04) is a simple seedling counter, and it is also available in local markets. It consists of a crystal oscillator and both a transmitter and receiver for ultrasonic waves, as shown in Figure 6. When triggered, the transmitter releases eight bursts of a directed 40 kHz of ultrasonic waves and starts a timer.

The ultrasonic pulses move outward until they encounter the SS, as shown in Figure 7. The SS causes the ultrasonic pulses to be reflected directly towards the receiver. The ultrasonic receiver detects the reflected ultrasonic pulses and stops the timer. The ultrasonic burst speed is 340 m/s in air. The distance (d) between the SS and transmitter can be identified based on the number of counts made by the timer. The time (T), rate, and distance (TRD) measurement formulae are expressed in Equation (1).

$$d = C \times T \tag{1}$$

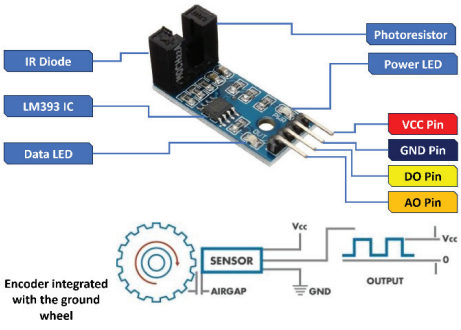


Figure 5. Main parts and connecting pins of the speed sensor used in the VRCMM. AO represents the analog pin, DO denotes the digital output pin, GND represents the ground pin, and VCC is the voltage common collector.

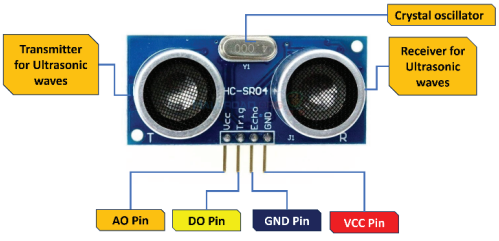


Figure 6. Main parts and connecting pins of the ultrasonic sensor used in the RSMS. AO represents the analog pin, DO denotes the digital output pin, GND represents the ground pin, and VCC is the voltage common collector.

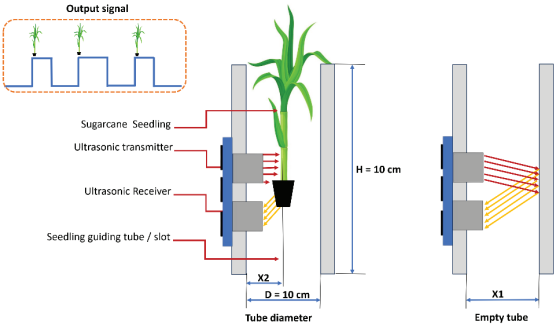


Figure 7. The main component and dimensions of the RSMS, where H is the height of SGS, D is the diameter of SGS, and X1 and X2 are the measured distances.

In this particular case, as it typically takes twice as long to record from transmitter to SS and then return to the receiver, the time is divided by two.

2.4. Variable-Rate Control Metering Mechanism (VRCMM)

The hardware components of the VRCMM were composed of four main electronic parts including an AMB, a speed sensor, stepper motor, and its driver, as shown in Figure 8.

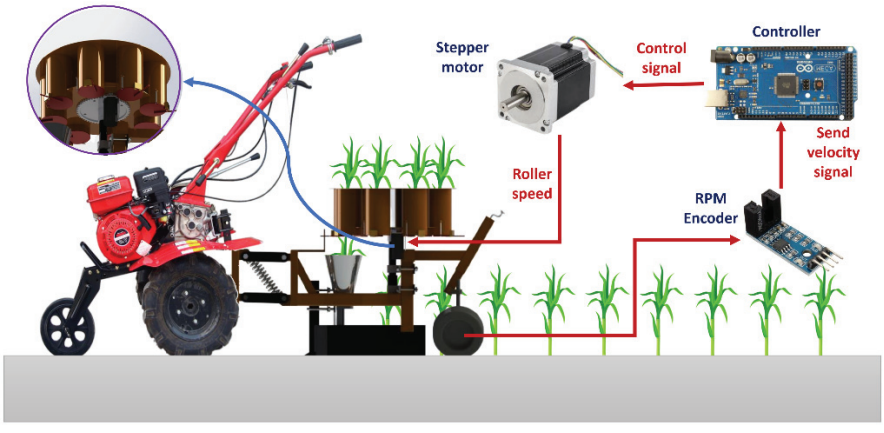


Figure 8. The schematic diagram of the VRCMM (Open-loop System).

2.4.1. Forward Speed

To design a precision sugarcane transplanting machine, the actual forward speed needs to be precisely measured. Where slippage occurred between the 2-wheel walking tractor and the soil, the speed sensor and encoder were connected to the ground wheel of the transplanter unit to identify its rotating and forward speed, as depicted in Figure 8. The ground wheel's RPM speed was detected, and the speed sensor supplied the output pulses to the AMB by a fixed time interval ΔT, where the travel speed v_f was calculated using Equation (2), according to Li et al. [49].

$$v_f = \frac{\pi \times d_2 \times C}{n \times \Delta T} \tag{2}$$

where d_2 is the ground wheel diameter in m, C denotes the quantity of pulses in ΔT, and n is the encoder resolution in pulses per revolution.

2.4.2. Operating Algorithm of the VRCMM

The operating algorithm of the VRCMM with an open-loop system is shown in Figure 9. The AMB (micro-controller) is the essential part of the complete control system. The data (signal) obtained by the speed sensor, the transplanting operation parameters (i.e., seedling type, seedling characteristics, appropriate speed ratio, interval space, ground wheel diameter, and slip ratio) set by Arduino IED software (version 2.2.1), and the programming code were uploaded to the AMB before using a USB cable. The stepper motor driver transmits the driving instructions of the stepper motor after uploading the required information for each component. The operating algorithm in Figure 9 is employed to read, control, and analyze the transplanting process. The open-loop algorithm was employed to manage the stepper motor speed according to the travel and forward speed, and the control system saved and transferred the transplanting and travel and forward speed to a laptop or smartphone application via a Bluetooth unit.

2.5. Seedling Counter System (SCS)

The SCS consisted of an SGS and a seedling cylinder as part of the seedling metering device. The SGS was developed to allow the falling seedlings to pass steadily through the ultrasonic waves. In addition, the seedling rotating tube can help seedlings to fall through the SGS smoothly, as shown in Figure 10.

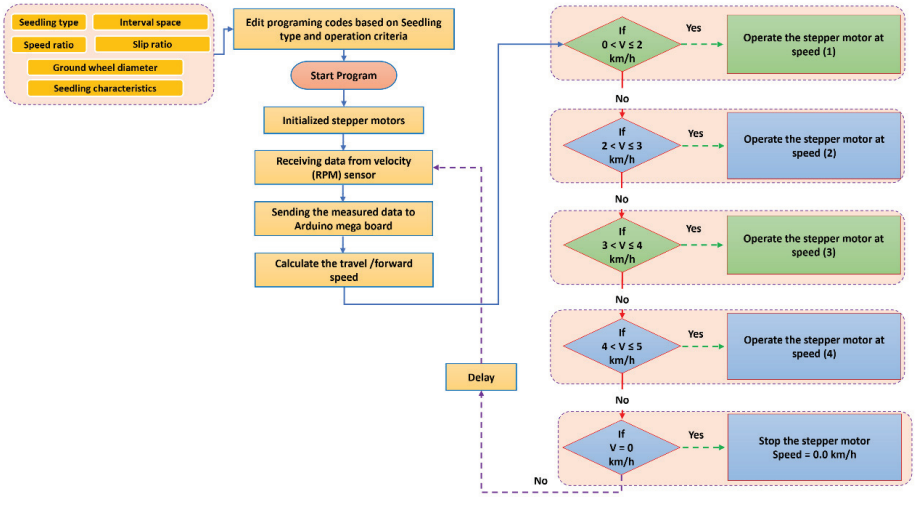


Figure 9. Operating algorithm flow chart of the VRCMM (open-loop system), where the first speed is 0–2 km/h, the second speed is 2–3 km/h, the third speed is 3–4 km/h, and the fourth speed is 4–5 km/h.

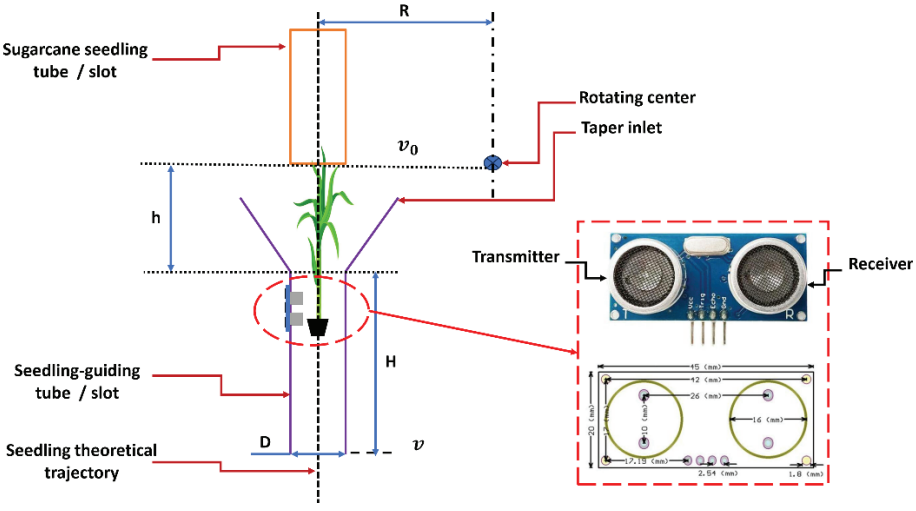


Figure 10. The falling trajectory when a seedling falls through the SGS, where R is the radius of the seedling cylinder, h is the vertical distance between the seedling cylinder and the taper inlet, H is the length of the SGS, D is the diameter of the SGS, v_0 is the initial velocity of the seedlings, and v is the final velocity of the seedling.

2.5.1. Calculation of SCS

As seen in Figure 11, when the SS was released by the seedling cylinder, its projected path was a parabola curve. The vertical distance between the inlet and the outlet of the SGS was 35 mm.

The rotational speed of the seedling metering mechanism can be altered by adjusting the pulse-width modulation (PWM) of the stepper motor speed controller. The rotation speed of each seedling metering device matches to each combination of transplanting speed and spacing. The rotation speed of the seedling metering mechanism can be calculated using Equation (3), according to Xie et al. [50].

$$N_m = \frac{5000 \times v_f}{3 \times x_p \times n_c} \tag{3}$$

where N_m refers to the rotation speed of the seedling metering device, r/min, v_f refers to the forward speed of the transplanter, km/h, and the x_p represents the theoretical transplanting distance, cm.

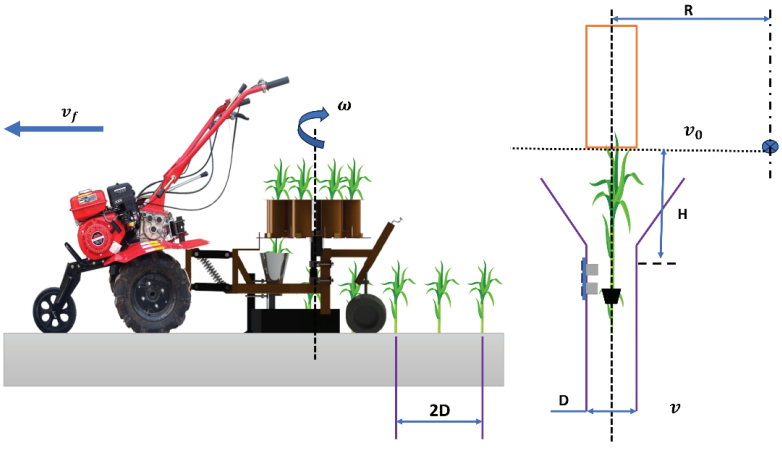


Figure 11. The falling trajectory when a seed falls through the SGS, where v_0 and v are the initial and final velocities of the seedling, respectively, v_f is the travel/forward speed, D is the transplant spacing, and ω is the rotating speed of the seedling metering device in rpm.

At the same time, the initial and final velocities of the SS can be calculated by using Equations (4)–(6) according to Xie et al. [50].

$$H = v_0 \times t + \frac{1}{2} g \times t^2 \tag{4}$$

where H is the height of seedling-guiding slot, m, v_0 is the initial speed of the seedling, m/s, and t is the time it takes for the seedling to reach the sensor from the seed outlet, s.

$$v = v_0 + g \times t \tag{5}$$

where $v_0 = 0$, v is the speed of the seedling passing the sensor, m/s, and g is the gravity acceleration = 9.81 m/s^2.

$$v = \sqrt{2g \times H} \tag{6}$$

2.5.2. Design of a Seedling Counter System (SCS)

The controller serves as a connection point between the seedling counter sensor and the software used for terminal display (on a laptop or smartphone), and it fundamentally performs the operating process as follows. Identifying and processing the electric signals required to have the target factors comes first. The next step is to achieve two-way real-time communication with the laptop or mobile phone and provide the transplanting factors to the terminal display software for real-time display, as well as the parameters that are returned by the monitoring terminal.

In this work, a laptop was utilized to develop a controller for the processor, which includes two main components: hardware parts and software program algorithms. The main job of the controller is illustrated in Figure 12. The process starts by acquiring the electrical signal created when the SS passes by the seedling counter sensor; the required parameters (number of seedlings, interval time between two consecutive seedlings, number of qualified seedlings, number of missed seedlings, qualified rate, and missed rate) are determined by the software program, which then assigns the parameter to the final variable

and sends it to the communication serial port according to the communication protocol. The Bluetooth communication module (HC-05) is connected the serial port, and the collected data are delivered there in real-time display by the laptop program. Figure 12 illustrates the operating algorithms for the seedling counter system; all the equations listed in the same figure were described by Tang at al. [51].

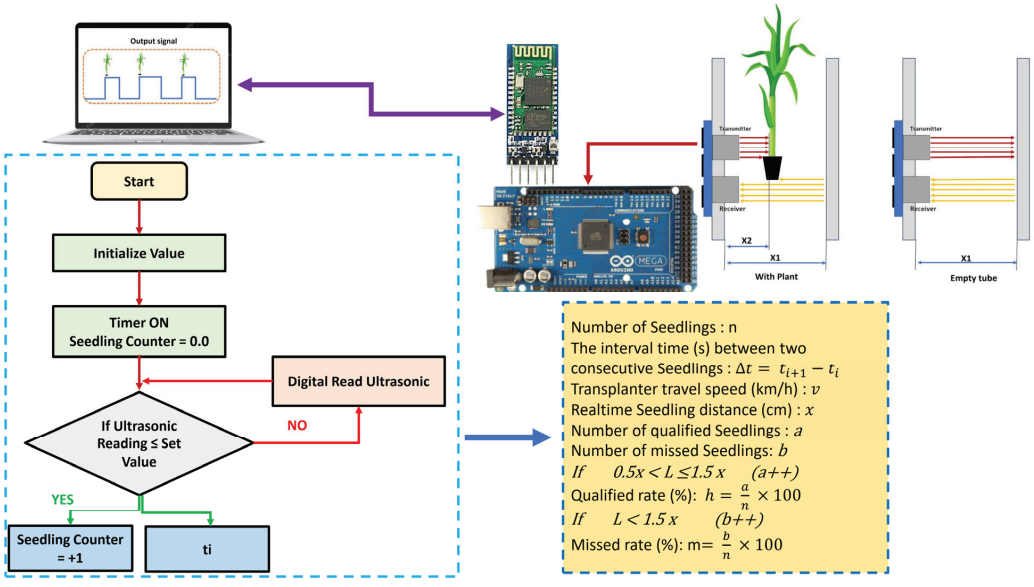


Figure 12. Operating algorithms for seedling flow rate detection, where *L* is the transplant spacing.

2.6. Remote Smart Monitoring System (RSMS)

In this investigation, the Arduino C programming languages, Mit-App inventor program, and Arduino IDE were used for the creation and design of the laptop application program. The laptop application can be used as seedling parameter monitoring software. It is capable of inputting the parameter values defined by the settings or the programming codes or display the monitored seedling factors in real time. In this instance, only one row of the transplanting parameter monitoring interface is employed because this is the seedling monitoring software for the single-row sugarcane transplanter. The software mainly comprises formulae like communication factor setting, transplanting parameter setting, and real-time transplanting factor display. The software's data update in real-time less frequently one in a second. Figure 13 shows the operating algorithm for the RSMS.

2.7. Laboratory Experiment

Comparative experiments were performed on a fabricated bench test to verify the effectiveness of the VRCMM and RSMS, at the Faculty of Agriculture and Natural Resources, University of Aswan, Egypt. In this manner, non-experimental variables such as slippage ratio and travel speed can properly be controlled, while the experimental levels can be altered efficiently. The Arduino Mega Controller was updated with the seedling information and programming codes. These data included the seedling type and characteristics, speed ratio, interval space, ground wheel diameter, slip ratio between the soil surfaces and the ground wheel, the theoretical seedling rate, the uploading time, the real seedling rate, and the regulating transplanting rate. Each experiment's operating time was determined by the time it took the virtual transplanter to complete a transplanting plot. The experimental data can be displayed on a laptop by a USB cable and saved in Excel files or displayed on the designed application only using serial communication software (Bluetooth). The examination bench was built of a seedling metering device, a stepper motor, a stepper

motor driver, a pinion and sprocket, a speed sensor, an AMB, and a seedling counter sensor. Furthermore, a taper inlet, an SGS, and metallic frames made up the seedling flow rate sensor.

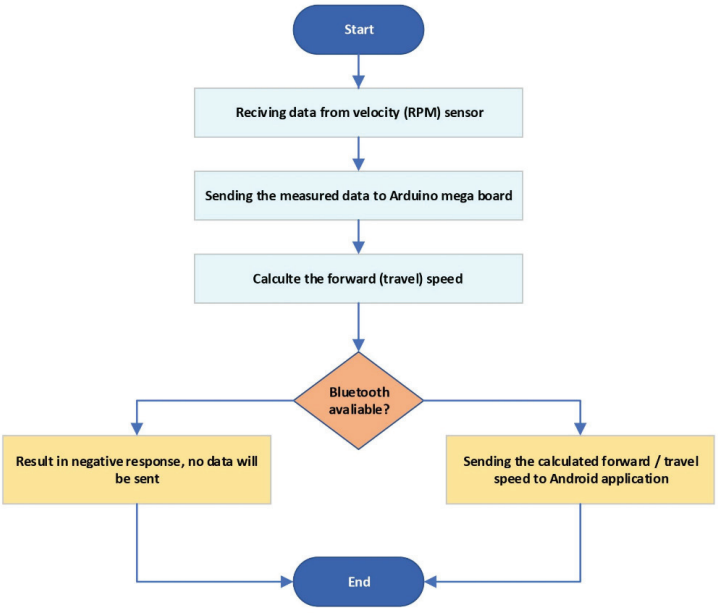


Figure 13. Operating algorithm for the RSMS.

2.8. Calibrating and Performance Evaluation of the VRCMM

The purpose of this test was to calibrate and assess the speed sensor's performance of the VRCMM under varying travel speeds and transplant spacing values. It aimed to find the rotating speed of the metering mechanism and the theoretical and actual transplant spacing under various seedling speeds. In addition, the test was run to verify the speed sensor accuracy of the seedling metering device at varying forward speeds and estimating relative error.

2.9. Performance Evaluation of the RSMS

This test examines the ultrasonic sensor's accuracy for monitoring seedling parameters at different travel speeds and transplant spacings. The goal was to find out the change law of the ultrasonic sensor's monitoring accuracy of single seedling flow at varying travel speeds and transplant spacings. It was run to discover the enhancement of precision of the ultrasonic sensor under the case of high travel speed and short transplant spacing. The test design is shown in Table 1. There was a total of twenty combinations (5 travel speeds × 4 transplant spacings), each of which was replicated three times, and each test had eight SSs.

Table 1. RSMS test design.

Speed Level	1	2	3	4	5
Travel speed (km/h)	2	3	4	5	6
Transplant spacing (cm)	20	25	30	40	

2.10. Statistical Analysis

A few statistical parameters were used, including coefficient of determination (R^2), to assess the performance of the speed sensor used in measuring processes of the travel

speed. Comparisons were performed among the data gathered from the speed sensor and the relevant (reference) data recorded by the stepper motor. R^2 denotes how much of the relationship between the measured and observed data is explained. Coefficient of determination (R^2) was computed using Equation (7), following the formula by Malone et al. [52].

$$R^2 = \frac{\sum(X - Y)^2}{\sum(X - Y_{ave})^2} \quad (7)$$

All parameters are explicated as follows: X is the independent value, Y is the dependent value, and Y_{ave} is the average value.

The relative error (E_R) is the ratio of the absolute error to the target or desirable speed, where absolute error is the difference between actual or measured speed and the target or desirable speed. By employing this statistical technique, we may estimate the magnitude of the absolute error relative to of the real measurement size. It provides an indication clue of how accurately the measurement is related to the dimension of the object being measured, which can be computed using Equation (8).

$$E_R = \frac{E_A}{V_T} \times 100 = \frac{|V_a - V_T|}{V_T} \times 100 \quad (8)$$

where E_A is the absolute error, V_a is the target speed, V_a is the actual or measured speed, km/h, and V_T is the target or desirable speed, km/h.

The coefficient of variance (CV) was determined to assess the seedling accuracy and seedling consistency of different experimental techniques. The computational methods of the abovementioned indicators are described in Equation (9), according to Liu et al. [53]:

$$CV = \frac{SD}{S_{ave}} \times 100 \quad (9)$$

where SD is the standard deviation and S_{ave} is the mean value.

Using the statistical software IBM SPSS Statistics version 25 and Microsoft Office (Excel 365, Version 2016), one-way analysis of variance (ANOVA) was run on the collected data at a 0.05 significance level.

3. Results and Discussion

3.1. Testing of the VRCMM

For validation in a precision sugarcane transplanter, the newly developed electric-driven control metering mechanism and smart monitoring system were tested under laboratory conditions. The stepper motor driver was operated based on the received signals from the AMB, which are a function of the travel speed. The speed sensor's output signals have to be calibrated by using the standard instrument to ensure the quality of data acquisition.

Before deploying the speed sensor for the VRCMM, the sensor was calibrated against the real speed of the stepper motor integrated with the ground wheel. The calibrations of the sensors were performed for the speed sensor (model: HC-89 RPM), as illustrated in Figure 14. The travel speed is plotted on the x-axis, and for every calibration condition, the speed recorded by the speed sensor was plotted on the y-axis.

Figure 14 shows the validation of the speed sensor measurements. With high R^2 values of 0.9986, the measured speed by the speed sensor has been validated with the observed measures by the stepper motor attached to the ground wheel, demonstrating an ideal fit between the travel speed and the measured speed by the speed sensor. At various forward speeds, the speed sensor performed well, and it had strong linear regressions (y = 1.001x) that nearly overlapped the 1:1 line.

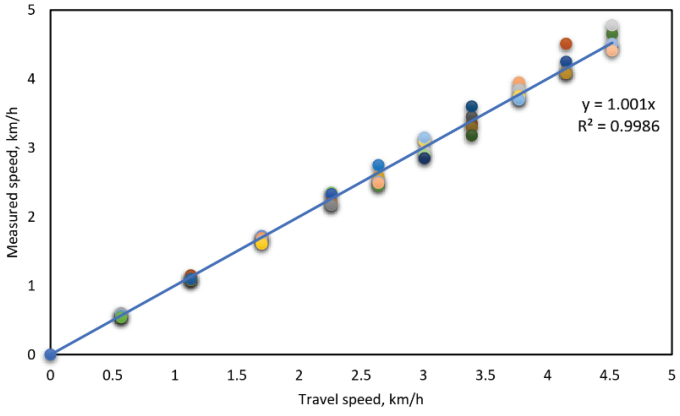


Figure 14. Tests on the speed regulation accuracy of the speed sensor of the VRCMM.

The accuracy of the stepper motor of the VRCMM's control system has a direct impact on the transplanting process. In the field, the transplanting process is controlled not only by the control system but also through the VRCMM's performance, the type of soil, soil state, slide between wheels and ground, etc. However, the bench test was mainly performed to detect the stability of the control system to keep the negative impact of the control system on the transplanting process at a minimum. As shown in Figure 15, the bench test was performed at 30 cm transplant spacing, and the travel speed of the ground wheel was selected between 0.5 km/h and 5 km/h, respectively. As depicted in Figure 16, the average relative error between the specified value of the VRCMM's stepper motor speed and the real value was 3.39%, and it increased with increasing the travel speed. The results further revealed that speed regulation was related with the transplanting index as reported by Che et al. [54], where they constructed a through-beam sensor with infrared sensing technology. The test findings revealed that the monitoring accuracy's relative error was between 0.5% and 5%.

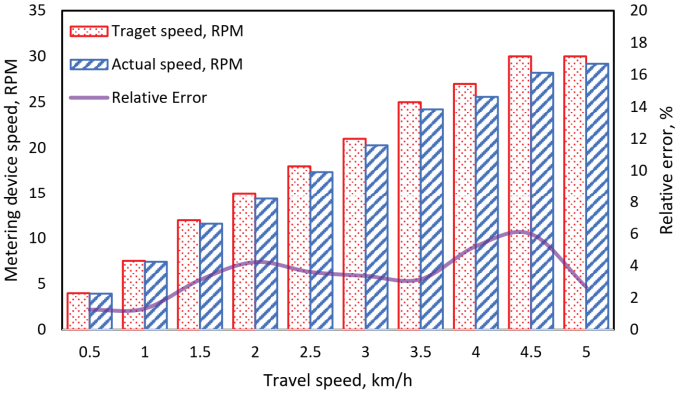
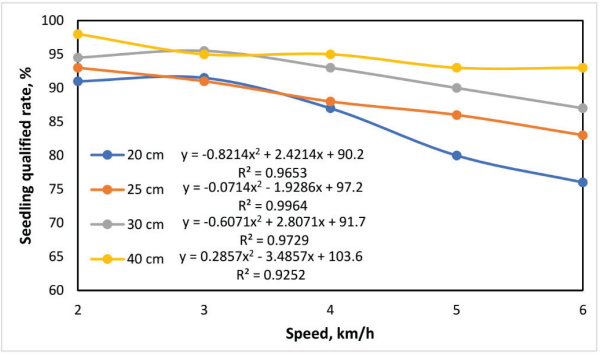
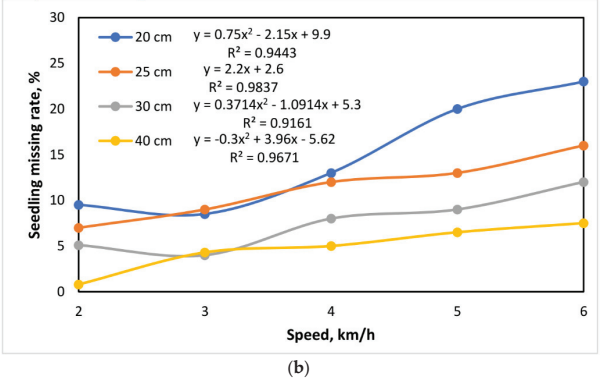


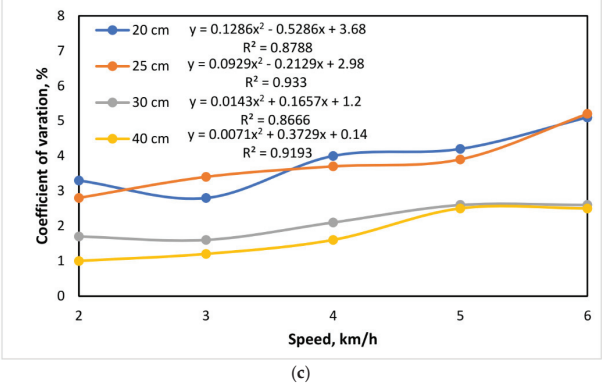
Figure 15. Speed accuracy detection of the VRCMM on bench test.



(**a**)



(**b**)



(**c**)

Figure 16. (**a**) Coefficient of determination of the quadratic relationships between travel speed on the RSMS and seedling qualified rate at different transplant spacings. (**b**) Coefficient of determin-tion of the quadratic relationships between travel speed on the RSMS and seedling missing rate at different transplant spacings. (**c**) Coefficient of determination of the quadratic relationships between travel speed on the RSMS and coefficient of variation at different transplant spacings.

3.2. The Effect of Travel Speed on the RSMS Results

Cay et al. [55] and Cay et al. [56] developed an electro-mechanic drive system (EMDS) for seed metering units. The findings of the field studies revealed that the speed of the seed plate is among the most crucial factors influencing the consistency of the planter's seed spacing. To estimate the impact of the travel speed on the obtained findings of the RSMS, the regression analysis of the travel speed was successfully performed under laboratory

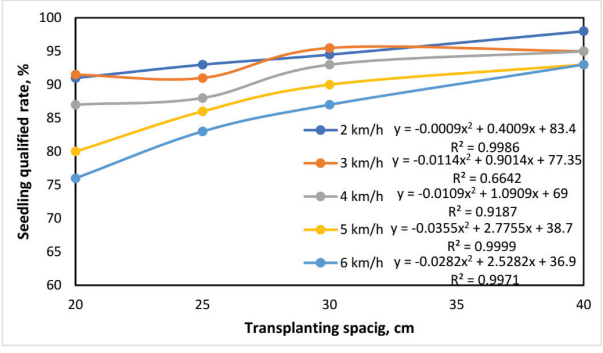
conditions at different transplanting spacings (20, 30, 40, and 50 cm) between seedlings at the same row. According to the preliminary analysis of the bench test outcomes, the accuracy of seedling counter sensor monitoring is least affected by travel speed within the range of 2–4 km/h. The bench test findings are relatively consistent, and also travel speed has remarkably affected the accuracy, in the range of 4–6 km/h. The obtained data from the bench test showed that travel speed has a greater impact on RSMS accuracy. As a result, the essential record of 4 km/h is chosen to execute the segmented regression analysis. Figure 16 demonstrates the regression analysis of the seedling counter sensor on three evaluation parameters at different transplanting spacings. Furthermore, Figure 16a shows that the regression fitting grade of the travel speed varies with transplant spacing. Overall, it is possible to assume that the R^2 is the same in terms of order, reaching R^2 20 cm > R^2 25 cm > R^2 30 cm > R^2 35 cm. Furthermore, the regression fit is best when the transplant spacing is 25 cm. This proves that when the transplant spacing is short, the change in travel speed has a significant impact on the RSMS results. The RSMS accuracy of the seedling counter sensor decreases as the travel speed increases. This is explicable by the fact that, when the travel speed increases, so does the rotation speed of the seedling metering device of the VRCMM increases, as well as an increase in the number of seedlings released per time period. The seedling counter sensor's RSMS accuracy decreases as the travelling speed of the seedlings increases concurrently. Broadly, the RSMS accuracy of the seedling counter sensor is remarkably affected when the travel speed is high.

Xie et al. [44] reported that the remote tracking system's monitoring performance is greater at slower speeds, and its monitoring error increases as the speed increases, and the speed has an effect on its precision. According to Xie et al. [50], when the seeding speed becomes high, the sensor monitoring reliability changes dramatically. Wu et al. [57] created a portable seed measuring system based on a programmable controller (PLC) and a touchscreen. The test results show that the monitored seed metering system provides the greatest seeding effect when the planter's forward speed is set to 5.23 km/h. Yang et al. [58] found that as the speed increases, the qualified index decreases, and the missed seed index increases. According to Karimi et al. [59], a faster seeding speed will lower the sensor's accuracy to monitor process.

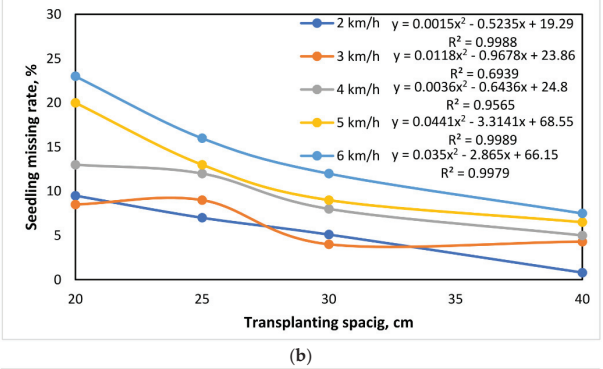
3.3. The Effect of Transplant Spacing on the RSMS Results

To assess the impact of transplant spacing on RSMS monitoring resultant outcomes, a transplant spacing regression analysis was performed at different travel speeds (2, 3, 4, 5, and 6 km/h). Similarly, preliminary examination of bench test results shows that in the 20–25 cm range, the bench test results change little, and transplant spacing has a bigger impact on RSMS accuracy; in the 25–40 cm range, the transplant spacing has a greater impact on RSMS accuracy. The bench test findings are reasonably steady, and transplant spacing has a significant impact on the seedling counter sensor's RSMS accuracy. As a result, the crucial point for segmented regression analysis was set at 25 cm transplant spacing. Figure 17 shows the regression analysis of the seedling counter sensor based on three criteria for assessment. Figure 17 makes it clear that the three assessment parameters of the seedling counter sensor have a similar law of change for transplant spacing, falling between 20 and 25 cm. Within the 25–40 cm range, there is a significant variance in the RSMS results; elsewhere, the changes are very consistent. This demonstrates that the RSMS accuracy is inconsistent and unsatisfactory in the 20–25 cm transplant spacing range. The RSMS accuracy is steady and good for transplant spacings between 25 and 40 cm. Furthermore, Figure 17 shows that the transplant spacing's degree of regression fitting varies with trip speeds. Overall, it can be concluded that the R^2 basically remains the same for all size orders: R^2 2 km/h, R^2 3 km/h, R^2 4 km/h, R^2 5 km/h, and R^2 6 km/h. Additionally, the regression fit is the best when the travel speed is 2 km/h. This proves that the change in transplant spacing has a greater impact on the RSMS results when the travel speed is low. The RSMS accuracy of the seedling counter sensor decreased as the transplant spacing fell, which may be a result of the fixed number of cylinders in the seedling cylinder.

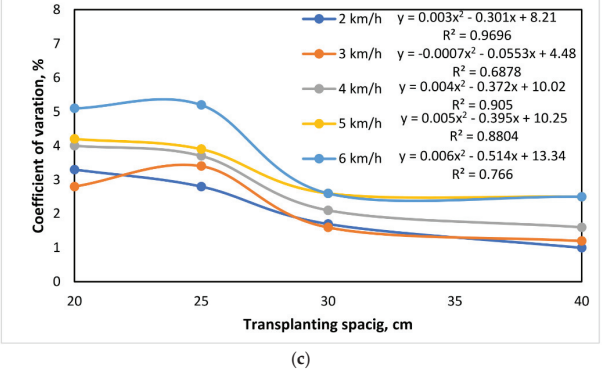
The speed of the seedling metering mechanism of the VRCMM can only be raised if the transplant spacing decreases to discharge more seedlings, and the number of seedlings discharged per unit time increases. The moving speed of the seedlings elevates in the meantime, which results in a decrease in the RSMS performance of the seedling counter sensor. In the case of small transplant spacing, the RSMS performance of the seedling counter sensor is impacted.



(**a**)



(**b**)



(**c**)

Figure 17. (**a**) Coefficient of determination of the quadratic relationships between transplant spacing on the RSMS and seedling qualified rate at different travel speeds. (**b**) Coefficient of determination of the quadratic relationships between transplant spacing on the RSM and seedling missing rate at different travel speeds. (**c**) Coefficient of determination of the quadratic relationships between transplant spacing on the RSM and coefficient of variation at different travel speeds.

Xie et al. [50] discovered that in cases where the seeding spacing is small, the sensor monitoring reliability changes dramatically. Further investigation revealed that the frequency of seeds going through the sensor is the most direct factor affecting sensor monitoring performance. The sensor monitoring reliability slowly decreases as frequency increases. Yang et al. [58] stated that speed has a significant impact on seeding performance, indicating that when speed increases, grain spacing dispersion becomes more extreme, the qualified index lowers, and the missing seed index rises.

4. Conclusions and Future Work

The potential of designing an automated sugarcane transplanter was investigated in this research. The study aimed to design and test the accuracy of a variable-rate control metering mechanism (VRCMM) and a remote smart monitoring system (RSMS) for a precision sugarcane transplanter based on IoT technology. The findings of the laboratory tests demonstrated the following:

- The speed sensor calibration results showed good performance, indicating perfect agreement between the travel speed and the speed measured by the speed sensor at different forward speeds during the experiment.
- The RSMS results are more affected by a change in travel speed when the transplant spacing is small. The RSMS accuracy of the seedling counter sensor decreased as the travel speed increased.
- When the travel speed is low, the change in transplant spacing has a greater impact on the RSMS outcomes. The RSMS accuracy of the seedling counter sensor decreased as the transplant spacing decreased.
- In conclusion, the obtained results of the current study are of great help to the design and improvement of a cost-effective, smart, and intelligent sugarcane transplanter.

Author Contributions: Conceptualization, A.E.E.; methodology, A.E.E., S.E., A.H.E. and A.H.M.; software, A.E.E., S.E., A.H.E., A.M.O. and Y.S.A.M.; validation, A.E.E., A.M.O., S.E., K.A.M., W.A.M., A.H.E. and Y.S.A.M.; formal analysis, A.E.E., A.M.O., A.H.M., K.A.M., W.A.M., S.E. and Y.S.A.M.; investigation, S.E., A.H.E. and A.H.M.; resources, A.E.E. and Y.S.A.M.; data curation, A.S.E., S.E., A.H.E., A.M.O., A.H.M., K.A.M., W.A.M. and Y.S.A.M.; writing—original draft preparation, A.E.E., S.E. and A.H.E.; writing—review and editing, A.E.E., Y.S.A.M., A.S.E., S.E., A.H.E., A.M.O., A.H.M., K.A.M., W.A.M. and visualization, A.E.E., S.E., A.H.E. and A.M.O.; supervision, A.H.E. and Y.S.A.M.; project administration, Y.S.A.M. and A.H.M.; funding acquisition, Y.S.A.M. All authors have read and agreed to the published version of the manuscript.

Funding: King Khalid University for funding this work through a large group Research Project under grant number (RGP 2/386/44).

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: All data are presented within the article.

Acknowledgments: The authors would like to acknowledge the Deanship of Scientific Research at King Khalid University for funding this work through a large group Research Project under grant number (RGP 2/386/44).

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Zein El-den, A.M.; Ahmed, S.F.; Hanafy, W.M.; Elwakeel, A.E. Review of some parameters related to the base-cutter of sugarcane harvesters. *Misr J. Agric. Eng.* **2020**, *37*, 325–330. [CrossRef]
2. Elwakeel, A.E.; Ahmed, S.F.; Zein Eldin, A.M.; Hanafy, W.M. A review on sugarcane harvesting technology. *Al-Azhar J. Agric. Eng.* **2021**, *2*, 54–63. [CrossRef]
3. Zein El-den, A.M.; Ahmed, S.F.; Hanafy, W.M.; Elwakeel, A.E. Fabrication and test of a tractor-front-mounted two-row sugarcane harvester. *Misr J. Agric. Eng.* **2020**, *37*, 331–344. [CrossRef]

4. Elwakeel, A.E.; Ahmed, S.F.; Zein Eldin, A.M.; Hanafy, W.M. Design and field testing of a sugarcane cutter. *Al-Azhar J. Agric. Eng.* **2021**, *1*, 39–48. [CrossRef]
5. Elwakeel, A.E.; Zein Eldin, A.M.; Tantawy, A.A.; Mohamed, S.M.A.; Mohamed, H.A.A. Manufacturing and performance evaluation of a sugarcane node cutting machine. *J. Soil Sci. Agric. Eng.* **2021**, *12*, 743–748. [CrossRef]
6. Wang, M.; Liu, Q.; Ou, Y.; Zou, X. Experimental study of the planting uniformity of sugarcane single-bud billet planters. *Agriculture* **2022**, *12*, 908. [CrossRef]
7. Singh, A.K.; Singh, P.R. Development of a tractor operated sugarcane cutter planter for mechanization of sugarcane planting in deep furrows. *Sugar Tech* **2017**, *19*, 416–423. [CrossRef]
8. Staggenborg, S.A.; Taylor, R.K.; Maddux, L.D. Effect of planter speed and seed firmers on corn stand establishment. *Appl. Eng. Agric.* **2004**, *20*, 573–580. [CrossRef]
9. Hossen, M.A.; Talukder, M.R.; Al Mamun, M.R.; Rahaman, H.; Paul, S.; Rahman, M.M.; Miaruddin, M.; Ali, M.A.; Islam, M.N. Mechanization status, promotional activities and government strategies of Thailand and Vietnam in comparison to Bangladesh. *AgriEngineering* **2020**, *2*, 489–510. [CrossRef]
10. Rípoli, M.L.C.; Rípoli, T.C.C.; Casagrandi, D.V.; Ide, B.Y. Evaluation of five sugar cane planters. In Proceedings of the 2007 ASAE Annual Meeting, Moscow, ID, USA, 13–15 September 2017; American Society of Agricultural and Biological Engineers: St. Joseph, MI, USA, 2007; p. 1.
11. Khedkar, M.B.; Kamble, A.K. Evaluation of mechanized planting of sugarcane. *Int. J. Agric. Eng.* **2008**, *1*, 136–139.
12. Tsuga, K. Development of fully automatic vegetable transplanter. *Jpn. Agric. Res. Q.* **2000**, *34*, 21–28.
13. Mazzetto, F.; Calcante, A. Highly automated vine cutting transplanter based on DGNSS-RTK technology integrated with hydraulic devices. *Comput. Electron. Agric.* **2011**, *79*, 20–29. [CrossRef]
14. Zhang, Q.A.; Wei, C.X.; Fan, X.H.; Shi, F.F. Chemical compositions and antioxidant capacity of by-products generated during the apricot kernels processing. *CyTA J. Food* **2018**, *16*, 422–428. [CrossRef]
15. Kumar, P.; Raheman, H. Automatic feeding mechanism of a vegetable transplanter. *Int. J. Agric. Biol. Eng.* **2012**, *5*, 20–27.
16. Shibusawa, S. Precision Farming and Terramechanics. In Proceedings of the Fifth ISTVS Asia-Pacific Regional Conference in Korea, Seoul, Republic of Korea, 20–22 October 1998.
17. Elwakeel, A.E.; Mazrou, Y.S.A.; Tantawy, A.A.; Okasha, A.M.; Elmetwalli, A.H.; Elsayed, S.; Makhlouf, A.H. Designing, optimizing, and validating a low-cost, multi-purpose, automatic system-based RGB color sensor for sorting fruits. *Agriculture* **2023**, *13*, 1824. [CrossRef]
18. McBratney, A.; Whelan, B.; Ancev, T.; Bouma, J. Future directions of precision agriculture. *Precis. Agric.* **2005**, *6*, 7–23. [CrossRef]
19. Yazgi, A.; Degirmencioglu, A. Measurement of seed spacing uniformity performance of a precision metering unit as function of the number of holes on vacuum plate. *Measurement* **2014**, *56*, 128–135. [CrossRef]
20. Xing, H.; Wang, Z.; Luo, X.; He, S.; Zang, Y. Mechanism modeling and experimental analysis of seed throwing with rice pneumatic seed metering device with adjustable seeding rate. *Comput. Electron. Agric.* **2020**, *178*, 105697. [CrossRef]
21. Karimi, H.; Navid, H.; Besharati, B.; Behfar, H.; Eskandari, I. A practical approach to comparative design of non-contact sensing techniques for seed flow rate detection. *Comput. Electron. Agric.* **2017**, *142*, 165–172. [CrossRef]
22. Karayel, D.; Wiesehoff, M.; Özmerzi, A.; Müller, J. Laboratory measurement of seed drill seed spacing and velocity of fall of seeds using high-speed camera system. *Comput. Electron. Agric.* **2006**, *50*, 89–96. [CrossRef]
23. Cay, A.; Kocabiyik, H.; Karaaslan, B.; May, S.; Khurelbaatar, M. Development of an opto-electronic measurement system for planter laboratory tests. *Measurement* **2017**, *102*, 90–95. [CrossRef]
24. Liang, Z.; Zhang, D.; Yang, L.; Cui, T.; Hao, Y. Experimental study on motor driven pneumatic precision seed-metering device for maize. In Proceedings of the 2015 ASABE Annual International Meeting, New Orleans, LA, USA, 26–29 July 2015; American Society of Agricultural and Biological Engineers: St. Joseph, MI, USA, 2015; p. 1.
25. Miller, E.A.; Rascon, J.; Koller, A.; Porter, W.M.; Taylor, R.K.; Raun, W.R.B.; Kochenower, R. Evaluation of corn seed vacuum metering systems. In Proceedings of the 2012 ASABE Annual International Meeting, Dallas, TX, USA, 29 July–1 August 2012; American Society of Agricultural and Biological Engineers: St. Joseph, MI, USA, 2012; p. 1.
26. Zhan, Z.; Yaoming, L.; Jin, C.; Lizhang, X. Numerical analysis and laboratory testing of seed spacing uniformity performance for vacuum-cylinder precision seeder. *Biosyst. Eng.* **2010**, *106*, 344–351. [CrossRef]
27. Pérez-Ruiz, M.; Slaughter, D.C. Development of a precision 3-row synchronised transplanter. *Biosyst. Eng.* **2021**, *206*, 67–78. [CrossRef]
28. Barbieri, P.; Echarte, L.; Della, M.A.; Sadras, V.O.; Echeverria, H.; Andrade, F.H. Maize evapotranspiration and water-use efficiency in response to row spacing. *Agron. J.* **2012**, *104*, 939–944. [CrossRef]
29. Doerge, T.; Hall, T.; Gardner, D. New research confirms benefits of improved plant spacing in corn. *Crop Insights* **2002**, *12*, 1–5.
30. Morente, A.B.; Elorza, P.B.; Díaz-Ambrona, C.G.H.; Da Costa, G.D.; Griepentrog, H.W. Even-sowing pattern strategies for a low-input organic system in forage maize. *Agric. Eng. Int. CIGR J.* **2013**, *15*, 171–179.
31. Perez-Ruiz, M.; Slaughter, D.C.; Gliever, C.; Upadhyaya, S.K. Tractor-based Real-time Kinematic-Global Positioning System (RTK-GPS) guidance system for geospatial mapping of row crop transplant. *Biosyst. Eng.* **2012**, *111*, 64–71. [CrossRef]
32. Samborski, S.M.; Gozdowski, D.; Stępień, M.; Walsh, O.S.; Leszczyńska, E. On-farm evaluation of an active optical sensor performance for variable nitrogen application in winter wheat. *Eur. J. Agron.* **2016**, *74*, 56–67. [CrossRef]

33. Duncan, W.G. A Theory to explain the relationship between corn population and grain yield. *Crop Sci.* **1984**, *24*, 1141–1145. [CrossRef]
34. Lowenberg-DeBoer, J. Economics of variable rate planting for corn. In Proceedings of the Proceedings of the Fourth International Conference on Precision Agriculture, St. Paul, MN, USA, 19–22 July 1999; Wiley Online Library: Hoboken, NJ, USA, 1999; pp. 1643–1651.
35. Razavi, J.; Namjoo, M. Determination of forward speed effect on planting uniformity in a sugarcane billet planter. In Proceedings of the XVIIth World Congress of the International Commission of Agricultural and Biosystems Engineering (CIGR) Paper No CSBE001142, Quebec City, QC, Canada, 13–17 June 2010.
36. Taghinezhad, J.; Alimardani, R.; Jafari, A. Optimization cane traction output from hopper in full-automatic sugarcane planters by using response surface modeling and analytical hierarchy process. *Agric. Eng. Int. CIGR J.* **2013**, *15*, 138–147.
37. Saengprachatanarug, K.; Wongpichet, S.; Ueno, M.; Taira, E. Comparative discharge and precision index of a sugar cane billet planter. *Appl. Eng. Agric.* **2016**, *32*, 561–567.
38. Saengprachatanarug, K.; Chaloemthoi, C.; Kamwilaisak, K.; Kasemsiri, P.; Chaun-Udom, S.; Taira, E. Effect of metering device arrangement to discharge consistency of sugarcane billet planter. *Eng. Agric. Environ. Food* **2018**, *11*, 139–144. [CrossRef]
39. Namjoo, M.; Razavi, J.; Khani, A. Fabrication and evaluation of a metering device for a sugarcane billet planter. *Yuz. Yil Univ. J. Agric. Sci.* **2015**, *25*, 1–12.
40. Liye, Z.; Jian, X. Study on optoelectronic sensor for performance detection of a seedmeter. *Trans. Chin. Soc. Agric. Mach.* **2005**, *36*, 41–43.
41. Chen, J.; Bian, J.; Li, Y.; Zhao, Z.; Wang, J. Performance detection experiment of precision seed metering device based on high-speed camera system. *Trans. Chin. Soc. Agric. Eng.* **2009**, *25*, 90–95.
42. Li, M.; Ding, Y.; Liao, Q. Loss sowing detection in field of pneumatic precision metering device for rapeseed. *Trans. Chin. Soc. Agric. Eng.* **2010**, *26*, 27–31.
43. *GB/T 6973-2005*; Standardization Administration of the PR China and General Administration of Quality Supervision, Inspection and Quarantine of the PR China. Testing Methods of Single Seed Drills (Precision Drills). Standard C.N.: Beijing, China, 2005.
44. Xie, C.; Zhang, D.; Yang, L.; Cui, T.; Zhong, X.; Li, Y.; Ding, Y.; Ding, Z. Remote monitoring system for maize seeding parameters based on android and wireless communication. *Int. J. Agric. Biol. Eng.* **2020**, *13*, 159–165. [CrossRef]
45. Raheman, H.; Kumar, R. An embedded system for detecting seed flow in the delivery tube of a seed drill. In Proceedings of the International Conference on Advances in Chemical, Biological & Environmental Engineering (ACBEE), Singapore, 29–30 March 2015; pp. 236–241.
46. Al-Mallahi, A.A.; Kataoka, T. Application of fibre sensor in grain drill to estimate seed flow under field operational conditions. *Comput. Electron. Agric.* **2016**, *121*, 412–419. [CrossRef]
47. Navid, H.; Ebrahimian, S.; Gassemzadeh, H.R.; Mousavi, N.M.J. Laboratory evaluation of seed metering device using image processing method. *Aust. J. Agric. Eng.* **2011**, *2*, 1–4.
48. Wang, C.P.; He, R.Y. Performance detection of precision seed-metering device based on single chip microprocessor. *Sci. Technol. Eng.* **2011**, *11*, 299–302.
49. Li, Y.; He, X.; Tao, C.; Zhang, D.X.; Song, S.; Rui, Z.; Wang, M. Development of mechatronic driving system for seed meters equipped on conventional precision corn planter. *Int. J. Agric. Biol. Eng.* **2015**, *8*, 1–9.
50. Xie, C.; Zhang, D.; Yang, L.; Cui, T.; Yu, T.; Wang, D.; Xiao, T. Experimental analysis on the variation law of sensor monitoring accuracy under different seeding speed and seeding spacing. *Comput. Electron. Agric.* **2021**, *189*, 106369. [CrossRef]
51. Tang, H.; Xu, C.; Wang, Z.; Wang, Q.; Wang, J. Optimized design, monitoring system development and experiment for a long-belt finger-clip precision corn seed metering Device. *Front. Plant Sci.* **2022**, *13*, 814747. [CrossRef] [PubMed]
52. Malone, B.P.; Styc, Q.; Minasny, B.; McBratney, A.B. Digital soil mapping of soil carbon at the farm scale: A spatial downscaling approach in consideration of measured and uncertain data. *Geoderma* **2017**, *290*, 91–99. [CrossRef]
53. Liu, W.; Hu, J.; Zhao, X.; Pan, H.; Lakhiar, I.A.; Wang, W. Development and experimental analysis of an intelligent sensor for monitoring seed flow rate based on a seed flow reconstruction technique. *Comput. Electron. Agric.* **2019**, *164*, 104899. [CrossRef]
54. Che, Y.; Wei, L.G.; Liu, X.T.; Li, Z.L.; Wang, F.Z. Design and experiment of seeding quality infrared monitoring system for no-tillage seeder. *Trans. Chin. Soc. Agric. Eng.* **2017**, *33*, 11–16. (In Chinese)
55. Cay, A.; Kocabiyik, H.; May, S. Development of an electro-mechanic control system for seed-metering unit of single seed corn planters Part I: Design and laboratory simulation. *Comput. Electron. Agric.* **2018**, *144*, 71–79. [CrossRef]
56. Cay, A.; Kocabiyik, H.; May, S. Development of an electro-mechanic control system for seed-metering unit of single seed corn planters Part II: Field performance. *Comput. Electron. Agric.* **2018**, *145*, 11–17. [CrossRef]
57. Wu, N.; Lin, J.; Li, B.F.; Zhou, Y.M. Design and test on performance monitoring system of no-tillage planter seed-metering device. *Trans. Chin. Soc. Agric. Mach.* **2016**, *47*, 69–75. (In Chinese)

58. Yang, S.; Wang, X.; Gao, Y.Y.; Zhao, X.G.; Dou, H.J.; Zhao, C.J. Design and experiment of motor driving bus control system for corn vacuum seed meter. *Trans. Chin. Soc. Agric. Mach.* **2019**, *50*, 57–67. (In Chinese)
59. Karimi, H.; Navid, H.; Besharati, B.; Eskandari, I. Assessing an infrared-based seed drill monitoring system under feld operating conditions. *Comput. Electron. Agric.* **2019**, *162*, 543–551. [CrossRef]

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.

agriculture

MDPI

Article

Development of Boom Posture Adjustment and Control System for Wide Spray Boom

Jinyang Li [1,2,*], Zhenyu Nie [1,2], Yunfei Chen [1,2], Deqiang Ge [1,2] and Meiqing Li [1,2]

[1] College of Agricultural Engineering, Jiangsu University, Zhenjiang 212013, China; m15951893771@163.com (Z.N.); 17826079289@163.com (Y.C.); qdgdq@163.com (D.G.); lmqljy@163.com (M.L.)
[2] Key Laboratory of Modern Agricultural Equipment and Technology, Ministry of Education, Zhenjiang 212013, China
* Correspondence: by0817136@163.com

Abstract: To obtain a more consistent droplet distribution and reduce spray drift, it is necessary to keep the entire spray boom parallel to the crop canopy or ground and maintain a certain distance from the spray nozzles to the crop canopy or ground. A high-performance boom active control system was developed for boom trapezoid suspension. The hydraulic system and hardware circuit of the boom control system were designed based on analyzing the configuration of active trapezoid suspension. The mathematical models of valve-controlled hydraulic cylinders and active boom suspensions were developed. Step response and frequency domain response analysis of passive suspension were conducted by Simulink simulations, and then key parameters of the boom suspension and hydraulic system were determined. A feedforward proportion integration differentiation (FPID) control algorithm was proposed to improve the tracking performance. The designed control system was assembled on a 24 m boom with trapezoid suspension. The response characteristic of the active boom control system was tested by the step signal and the sinusoidal signal from a six-degree-of-freedom hydraulic motion platform. Firstly, the tracking performance of the active balance control system for the PID (proportion integration differentiation) and FPID control algorithms was compared for a given 0.2 Hz sine signal. Then, for the ground-following control system, the response characteristics in challenging terrain and tracking performance in less challenging terrain were tested. Field experiment results indicate that the maximum rolling angle of the chassis was 3.896° while the maximum inclination angle of the boom was 0.453°. The results show that the designed boom adjustment and control system can effectively adjust the boom motion in real time and meet the requirements of field operation.

Keywords: boom sprayer; active balance; ground following; motion control; feedforward proportion integration differentiation; vibration

Citation: Li, J.; Nie, Z.; Chen, Y.; Ge, D.; Li, M. Development of Boom Posture Adjustment and Control System for Wide Spray Boom. *Agriculture* **2023**, *13*, 2162. https://doi.org/10.3390/agriculture13112162

Academic Editors: Jin Yuan and Zichen Huang

Received: 7 September 2023
Revised: 9 November 2023
Accepted: 11 November 2023
Published: 17 November 2023



Copyright: © 2023 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (https://creativecommons.org/licenses/by/4.0/).

1. Introduction

The rapid development of land transfer in China promoted the emergence of large and medium-sized farms. A variety of large agricultural machinery with high operation efficiency is widely used in the agricultural industry. Due to advantages such as wide operation spraying swath, precise quantification of spraying, and uniform spraying amount, sprayers with above 24 m spraying swath are widely used in field crop protection. However, the large flexibility of the boom was caused by the wide spray boom. Since the ground surface is undulating in the field, the small vibration of the sprayer body will cause a large fluctuation of boom ends. The unexpected boom movements frequently happen as the sprayer travels over bumpy ground or undulating fields. This not only affects the uniformity of spray quantity distribution but also leads to a decline in application quality [1]. In severe cases, the boom ends will touch the ground or crop canopy, causing damage to the boom and sprayer and injury to the crop. To ensure working quality and

safety, the driver of the sprayer should observe whether there is contact between the boom and the crop during the operation and thus adjust the tilt angle of the boom manually, which increases the labor intensity of the driver.

The vertical rolling motion has a significant effect on spray distribution pattern [2,3] and droplet spray uniformity [4–6]. To obtain a more consistent droplet distribution and reduce spray drift, it is necessary to keep the spray boom parallel to the crop canopy or ground and maintain a certain distance from the spray nozzle tips to the crop canopy or ground. Unfortunately, the passive suspensions hardly always make boom maintain parallel with the crop canopy or ground due to the deflection of the vehicle tires and sinkage of the wheels into the ground. To make up for the deficiency of passive suspension, a variety of active suspensions were investigated [7–15].

Different control strategies, such as proportional-integral (PI), optimal control [16], state feedback [17], sliding mode control [18], and H∞ theory [19], were used to adjust the boom motions in the vertical and horizontal planes. To improve the response ability and tracking performance of the active boom suspension, the FPID (feedforward proportion integration differentiation) control algorithm is proposed by integrating the forward control link with the PID. As an open-loop control, the forward feedback control can directly exert the correction effect according to the disturbance quantity when the disturbance occurs. In theory, feedforward control can eliminate the deviation caused by disturbance. At present, feedforward control is widely used in electro-hydraulic servo systems and hydraulic systems. The velocity feedforward compensation proportional integral control (VFPID) was used to control the active pendulum boom suspension and effectively control the boom movement [13]. However, the forward feedback relies on a precise system model to eliminate the deviation caused by disturbance. Therefore, the mathematical model of the boom chassis was established to obtain rapid response performance.

Several automatic control devices were developed for spraying height control by manufacturers such as Muller, Norca, and Ravel. Satisfactory control performance was obtained in flat, large plains when the boom sprayer travels across the field at relatively low speed. However, few imported spray height controllers have been used by domestic sprayers. The main reasons were as follows: (1) The control performance was discounted when these sprayers passed over challenging terrain, such as hilly and mountainous cultivated land, so that the boom hardly kept the desired spraying height at the desired response speed. Unfortunately, most of the cultivated fields are situated in hilly and mountainous areas of China. It is necessary that the high response speed of the electro-hydraulic servo system built in active boom suspension be required to meet the challenge of the complex environment; (2) it is difficult to match these imported control devices with special requirements for machine structure and hydraulic-servo system design with boom sprayers made in China; (3) these imported control devices were expensive, and the farmers could not afford to use them. The UF02 mounted spray from Amazone and Albators from Lennken are more than 90,000 pounds; (4) The control strategy assembled in these devices was private so that the manufacturer's technicians are required to repair it in the event of a failure, which results in increased maintenance costs and non-working time of the sprayers.

Based on the above reasons, although most large-sized sprayers are equipped with boom suspension, only a few controllers [13] whose control performance needed to be further improved were developed by manufacturers.

To reduce costs and improve response performance in challenging environments, many different attempts were made. A high-performance boom control system for pendulum active suspension was designed based on a DSP microprocessor and velocity feedforward compensation proportional integral control algorithm [14], which can realize effective control for the boom movement and meet the requirements of field use in less challenging terrains. Although the proposed control system provided theory and references for relative research, its performance can be further improved by developing an accurate system model.

The AMAZONE ContourControl device was assembled on the UF2002-mounted sprayer, which could realize fully automatic boom guidance, including boom height, tilt

adjustment, and boom lifting on the headland. This allowed the boom to maintain an optimal distance to the target surface across the whole working swath, even on extremely hilly terrain [20]. The boom height and boom angling of the LEMKEN Primus trailed sprayer were easily adjusted by the advanced hydraulic control system, and the boom damage was avoided by the advanced Magneo balance adjustment mechanism [21]. However, these two sprayers were tested in a cotton field in Xinjiang, where there are unforeseeable dark pits to lose to 1 m in the cotton fields. The test results indicated the response speed of the electro-hydraulic servo system built in these two kinds of active boom suspension was discounted, resulting in frequent damage to the boom. The main reason was that the response speed of the electro-hydraulic servo system was not fast enough for the challenging terrain. In this study, to acquire high response performance, sensing one-way valves, high-speed proportional valves, and hydraulic cylinders were used in the design of the hydraulic servo system. Furthermore, the FPID control algorithm was proposed, which was able to eliminate disturbances before the controlled quantity variable occurred.

The object of this study is to develop an automatic control device with low-cost and high-response spaying height for trapezoidal active suspension, aiming to maintain the desired spraying height from the crop canopy or ground in challenging terrain. The detailed aims include: (1) realize the automatic balance of boom in flat terrains; (2) ensure boom automatic following with crop canopy or ground in sloping and undulating terrains with rapid response; (3) realize the quick raise of boom wing in challenging emergency circumstances such as unilateral tire of sprayer sinkage into the ground quickly and avoid damage because of touching between boom tip and ground.

2. Materials and Methods

2.1. Total Configuration of Active Trapezoid Suspension

The schematic diagram of active trapezoid suspension is shown in Figure 1.

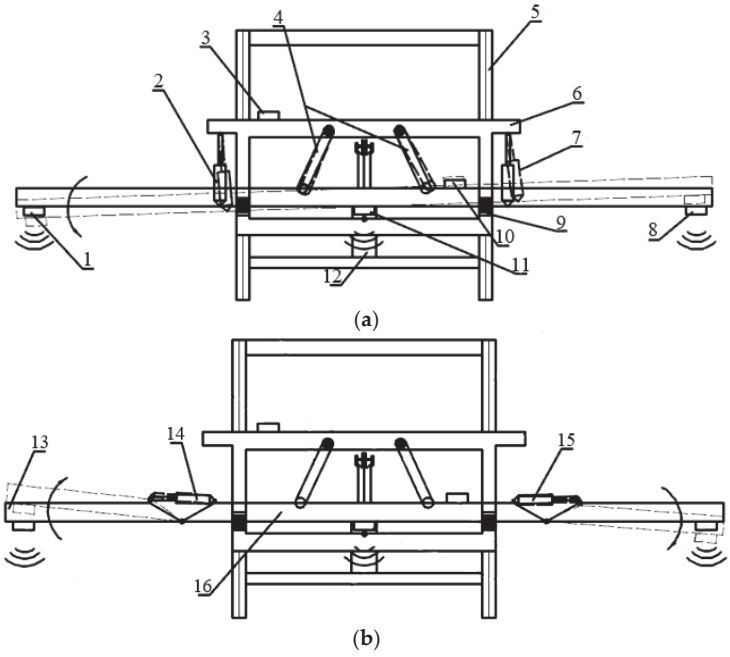
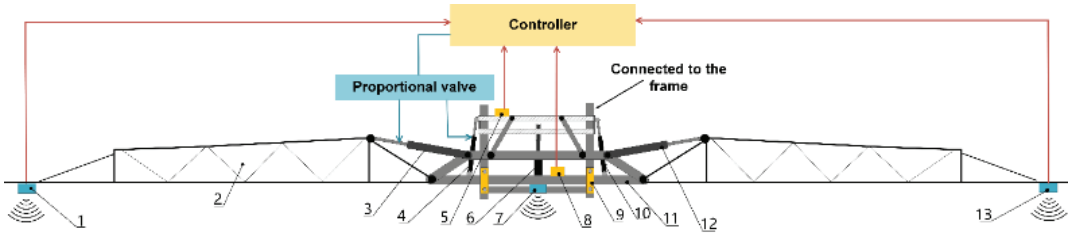


Figure 1. *Cont.*



(**c**)

Figure 1. Boom suspension geometry. (**a**) boom balance; (**b**) boom imitation; (**c**) boom balance and boom imitation: 1, 8, 12—Ultrasonic sensor; 2, 7—balance hydraulic cylinder; 3, 10—inclination angle sensor; 4—link rods of trapezoidal suspension; 5—boom suspension; 6—boom chassis; 9—mechanical limit device; 11—lift hydraulic cylinder; 13—sprayer boom; 14—left imitating hydraulic cylinder; 15—right imitating hydraulic cylinder; 16—center frame of boom.

The boom sprayer is mounted to the tractor by a three-point suspension mechanism, and the boom is attached to the sprayer via the boom chassis. The entire boom is divided into the left arm on the left side, the right arm on the right side, and the central frame situated in the middle part. The bottom of the left and right arms is hinged to the center frame, and the top parts of the left and right arms are articulated to the counterparts of the middle frame by the imitating hydraulic cylinder 14 or cylinder 15, respectively. Under the action of the following cylinders, the left arm and/or right arm can be revolved through the revolute joints to make unilateral or two-side boom arms that keep parallel to the ground or crop canopy. The boom lifting mechanism is comprised of a lifting hydraulic cylinder, lifting frame, fixed pulley, steel rope around the pulley, and a V-shape guide rail on the boom suspension frame, which is used to adjust the height of the entire boom. The mechanical limit device and polyurethane dampers are adopted to suppress the horizontal movements forward and backward and absorb the shock of the horizontal load. The two ends of the steel rope are fixed to the center frame. One end of the lifting hydraulic cylinder is welded to the boom suspension, and the other end is attached to the fixed pulley. As the lifting hydraulic cylinder retracts (extends), the whole boom ascends (descends). The boom suspension is used to bear the weight of the boom and the inertial load. The ends of balance hydraulic cylinder 2 and hydraulic cylinder 7 are jointed to the boom and boom lifting frame by revolute joints, respectively. The center frame rotates in a counterclockwise direction while the balance hydraulic cylinder 2 retracts (extends) and the balance hydraulic cylinder 7 extends (retracts). The maximum rolling angle of the boom relative to the chassis is determined by the mechanical limit device. Three ultrasonic sensors are adopted to sense the distance between the boom and the ground/crop canopy. Two inclination angle sensors (3 and 10) are mounted on the boom chassis and the center boom, which are used to collect the inclination angles of the boom chassis and the center boom, respectively.

2.2. Design of a Hydraulic System

Figure 2 shows the schematic diagram of the hydraulic system, which is composed of an active balance hydraulic circuit, ground imitation circuit, lifting hydraulic circuit of entire boom suspension, and oil supply circuit.

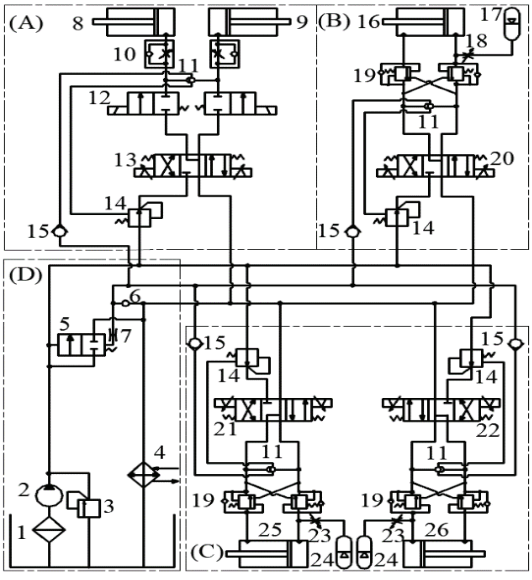


Figure 2. Schematic diagram of the hydraulic system. (**A**) active balance hydraulic circuit; (**B**) lifting hydraulic circuit of the entire boom suspension; (**C**) ground imitation circuit, and (**D**) oil supply circuit: 1—oil filter; 2—constant displacement pump; 3—overflow valve; 4—oil cooler; 5—constant difference overflow valve; 6—pressure relief opening; 7—damper 8, 9—active balance hydraulic cylinder; 10, 18, 23—one-way throttle valve; 11—shuttle valve; 12—two-position two-way solenoid valve; 13, 20, 21, 22—three-position four-way proportional valve; 14—pressure compensator; 15—sensing one-way valve; 16—lifting hydraulic cylinder; 17, 24—energy accumulator; 19—balance valve; 25, 26—imitation hydraulic cylinder.

Five cylinders, which are controlled by four three-position, four-way proportional valves, are adapted to control the boom attitude. Cylinders 8 and 9, which are controlled by valve 13, are used to keep the boom parallel to the ground or crop canopy. Under the action valve 13, the piston rod of cylinder 8 extends (retracts) and that of cylinder 9 retracts (extends), thus the whole boom is slanted to left or right and the boom balance is realized. Cylinders 25 and 26, which are controlled by valves 21 and 22, respectively, are adapted to rotate the left-side boom arm and right-side boom arm, respectively. Cylinders 25 and 26 can work along or together. Under the action of valve 21, the left-side boom arm is lifted (lowered) when the piston rod of cylinder 25 extends (retracts). The cylinder 16 is applied to adjust the height of the whole boom suspension, which is a one-way cylinder and controlled by proportional valve 20. The boom suspension is lifted when the piston rod extends. The lowering movement is accomplished by the gravity of the boom. Thus, the lifting and lowering of the whole spray boom is achieved. The balance valve 19 plays the role of locking the hydraulic oil road. The hydraulic system unloads, and the piston rod of the cylinder stops moving so that the boom keeps the current fixed. The overflow valves 3 and 5, pressure compensator 14, and sensing one-way valve 15 are used to maintain the pressure stability of the hydraulic pipelines.

2.3. Hardware Circuit of the Boom Control System

The hardware circuit of the boom control system is shown in Figure 3. To meet the real-time requirement of the control system and ensure system stability, the high-quality embedded industrial control module EMB8610I (Zhong Qian Ling Yun Electronic Co., Ltd., Beijing, China), whose MCU STM32F107VCT6 (STMicroelectronics, Co., Ltd., Agrate Brianza, Italy) was 32-bit ARM Cortex-M3 processor with 72 MHz main frequency was

selected as the controller. Moreover, EMB8610I integrated two RS-232, one RS-485, one CAN, one TTL Uart communication interface, two pulse-width modulation (PWM) output channels, seventeen 12-bit analog voltage/current input channels with a range of 0–10 V or 0–20 mA (switching voltage or current mode by jumper), and two 12-bit analog voltage output channels with a range of 0–10 V or 0–5 V (selecting by jumper).

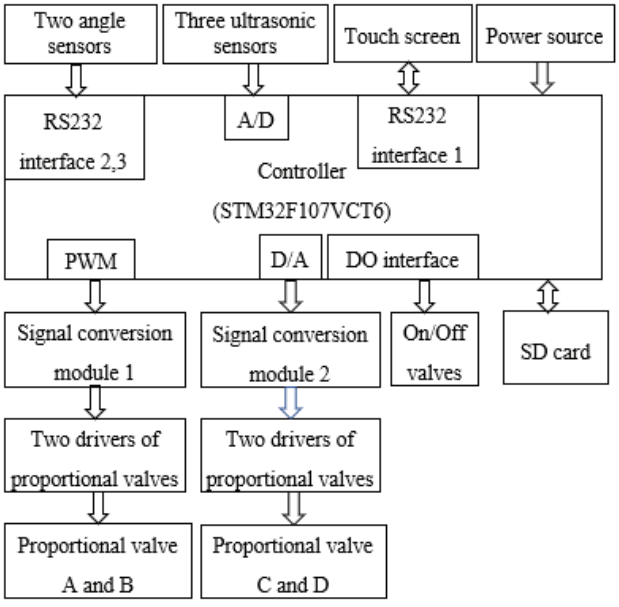


Figure 3. Schematic diagram of the controller.

The output signal of the ultrasonic sensor is an analog quantity in the range of 0–20 mA, so that three ultrasonic sensors can be connected directly to an analog voltage/current input channel. Similarly, the dynamic inclination angle sensor with digital output signals was attached to the controller by the RS-232 interface. The touch screen (DC10600M070_2111_0C, Dacai Optoelectric Co., Ltd., Guangzhou, China) with two RS-232 serial interfaces was adapted to set control parameters, transmit control commands, display the running state, and store sensor data for the control system. The communication between the touch screen and the controller was achieved by the RS-232 serial interface.

Three ultrasonic ranging sensors UC2000-30GM-IUR2-V15 (Pepperl+Fuchs, Co., Ltd., Bühl, Germany) were adopted as sensing the relative distance between the sprayer boom and crop canopy or ground. Digital drivers control the current to the solenoid of Atos proportional valves without a transducer, according to the electronic reference input signal. One E-BM-AS driver can drive two single or one double solenoid proportional valve. The solenoid proportionally transforms the current into a force, acting on the valve spool or poppet against a reacting spring, thus providing hydraulic regulation. The input voltage ranges from −10 V to +10 V. To achieve voltage matching, the linear conversion from D/A output voltage 0 V to 10 V to the input voltage of digital electronic drivers in the range of −10 V–+10 V was performed by the signal conversion module 2. Likewise, signal conversion module 1 was used to convert PWM voltage ranging from 0–3.3 V to −10 V–+10 V.

Two six-axis dynamic inclinometer/attitude sensors (LPMS-IG1-RS232, Alubi electronic technology Co., Ltd., Guangzhou, China) were used to measure the inclination angles of the sprayer boom and boom suspension, respectively.

2.4. Processing Method for Sensor Data

Irregular waggling occurred when the crop was subjected to wind, and thus signals from ultrasonic sensors were inevitably mingled with noise components. Moreover, the ultrasonic sensor signals were polluted by noise from the rough ground or heterogeneous crop canopy. To keep a pre-set distance between the boom and crop canopy or ground and to accomplish boom following undulating ground, accurate sensor data were the key and premise. Otherwise, the sprayer boom, under the action of the control system, will operate mistakenly. Therefore, to obtain effective and accurate sensor data, a combined filtering algorithm containing limiting filtering, median filtering, and moving average filtering (MAF) was adopted.

2.4.1. Moving Average Filter

To eliminate or reduce periodic interference and guarantee real-time sensor data, a moving-average filtering algorithm was adopted. The detailed expression is as follows:

$$HA_i = \frac{H_i + (i-1)HA_{i-1}}{i} \quad (1)$$

where H_i is the height value from an ultrasonic sensor at the current sampling cycle; i is the moving number and can be altered by touch screen. HA_{i-1} and HA_i are moving average height values from previous and current cycles, respectively.

2.4.2. Moving Median Filtering

A moving median filtering algorithm was used to suppress the impulse noise. The detailed algorithm can be seen in Ref. [14] and is not repeated here.

2.4.3. Limiting Filtering

Crop density at different growth stages was various. The sound reflection properties (intensity or rate) of the ultrasonic sensor on broad leaf crops such as cotton, soybeans, and potatoes and linear leaf crops such as wheat and rice were distinct.

At the seedling stage, the sound wave was mainly echoed by the ground because the growth density was relatively small. In this case, the ground feedback mode was adopted. With the development of growth, the rate of crop canopy to ground climbs so that the sound signal is mostly reflected by the crop canopy. Under the circumstances, crop height feedback mode was selected. However, in fact, the crop density and height were uneven and discontinuous, and a passage of area beneath the boom was in the absence of crop in crop height feedback mode, which caused impulse interference. At this time, the control system will mistakenly malfunction so that the boom vibrates and hardly keeps parallel to the crop canopy at the desired distance. To avoid the existence of the above phenomenon and improve the robustness of the boom control system, a limiting filtering algorithm [14] was added to the data processing of ultrasonic sensors. The crop average height (h_a) was manually measured at multiple points. The desired spraying height (h_s), h_a and the control parameters were set beforehand by touch screen prior to the application after the start of the control system. Meanwhile, the operation modes will be chosen.

2.5. Design of a Second-Order Low-Pass Filter Circuit for PWM Output

Each proportional valve needs a one-way analog signal to be driven. Four proportional valves were used in this study, and thus four-way individual analog signals were required. However, only two individual one-way analog signal output interfaces were available. Considering the cost and simplicity of the controller, double PWM output interfaces were adopted to simulate the D/A interface by adjusting the duty cycle to drive the other two proportional valves. To eliminate or reduce the effect of the first harmonic on PWM signals, a second-order low-pass filter circuit was designed. The determination of cut-off frequency and filter parameters was as follows:

The influence of the first harmonic on the output voltage should not exceed the accuracy of one bit. Under 12-bit resolution, the accuracy of one bit is $3.3\ /2^{12} = 0.0008$ V. Supposing $VH = 3.3$ V and $VL = 0$ V, the maximum of fundamental harmonic is $3.3 \times \frac{2}{\pi} = 2.10191$ V, the minimum attenuation value is at least $-20\lg\left(\frac{2.10191}{0.1289}\right) = -44$ dB.

Because the highest frequency Hsysclk of system clock is 72 MHz, the PWM frequency (fundamental harmonic frequency, f_1) is 281.25 kHz ($f_1 = Hsysclk/2^8 = 281.25$ kHz). For first-order RC filtering, the required cut-off frequency, $f_{p1} = 1.774$ kHz) can be calculated by the formula: $-10\lg\left[1 + \left(f_1/f_{p1}\right)^2\right] = -44$ Db. Likewise, the required cut-off frequency ($f_{p2} = 24.26$ kHz) for second-order RC filtering can be calculated using the formula: $-20\lg\left[1 + \left(f_1/f_{p2}\right)^2\right] = -44$ dB.

The amplitude-frequency characteristic of first-order RC filtering is: $-10\lg\left[1 + \left(f_1/f_{p1}\right)^2\right]$.

2.6. Experimental Design

2.6.1. Development of a Laboratory Test Bench

To validate the effectiveness of the proposed spray boom control system, various tests in the laboratory were performed under diverse working conditions through a six-degree-of-freedom hydraulic motion platform. The program of the boom controller, including the sensor data acquisition algorithm, signal processing algorithm, FPID control algorithm, and control instruction output, is developed in the C language. The sampling cycle of the control system is 10 ms.

Two posture sensors (LPMS-IG1, Alubi, Co., Ltd., Guangzhou, China) are adopted to sense the inclination angles of the center boom and support chassis, whose angle resolution is 0.01°, and the output frequency is 100 Hz. Three ultrasonic sensors (UC2000-30GM-IUR2-V15, Pepperl+Fuchs, Co., Ltd., Bühl, Germany) are used, whose maximum detection distance is 2000 mm and repetition accuracy is less than 0.1% of full scale. The specific installation positions are as shown in Figure 1.

Four proportional valves (DHZO-A-073-L3, Atos Co., Ltd., Sesto Calende, Italy) were selected to drive the left-imitating cylinder, right-imitating cylinder, lifting cylinder, and active balance cylinders, respectively. To improve the drive capability of the analog signal from the controller, four digital electronic drivers (E-BM-AS-PS-05H 12, Atos Co., Ltd., Sesto Calende, Italy) were selected to drive the proportional valves.

The characteristics of a hydraulic power unit are: a supply pressure of 12 Mpa and a maximum flow rate of 45 dm^3/min.

The developed control system is integrated the trapezoid suspension of a 24 m spray boom. The boom suspension geometry is shown in Figure 1. A test bench for the active boom suspension is shown in Figure 4.

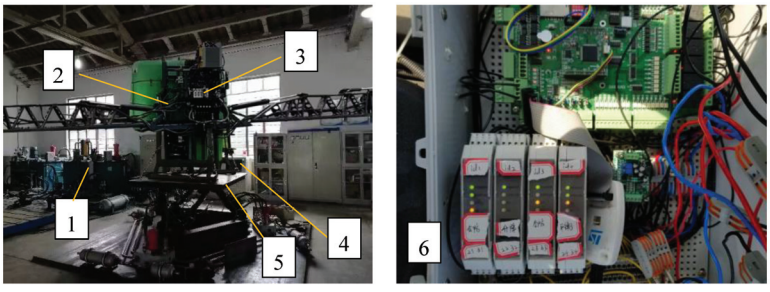


Figure 4. Test bench of the boom control system. 1—hydraulic power unit; 2—boom and trapezoid suspension; 3—control valves; 4—control computer of hydraulic test bench; 5—six degrees freedom motion planform; 6—boom controller.

Firstly, the experiments were carried out based on three various control modes (active balance mode, ground following mode, hybrid mode of active balance and ground following) of the control system. The active balance mode is used only when the difference between the boom angle and the chassis is greater than the set threshold. The ground-following mode is adopted when the absolute difference between the setting height and the actual height of the boom end is greater than the set threshold. The hybrid mode of active balance, and ground-following mode is a combination of the above two modes.

Secondly, the response characteristic of the control system for steep and abrupt undulations from the field ground for sinkage of the wheels into the ground was tested. Each run of the test is repeated three times, and the data of interest is averaged, which can be adopted for subsequent data analysis. The six-degree-of-freedom hydraulic motion platform is mainly used to simulate undulate terrains by providing a low-frequency variation signal with various amplitudes. Before testing, the boom is fixed on the hydraulic motion platform, which implies that the boom and hydraulic motion platform have the same motion mode.

During the process of testing, the inclination angles of the boom and boom suspensions, as well as the boom end heights, are collected in real time by the corresponding sensors integrated into boom. Meanwhile, the collected data are stored in the SD card for subsequent analysis. Based on the obtained data and the set values, the proposed control algorithm is conducted, and the valves and cylinders act. Thus, the boom posture is adjusted by the cylinders.

2.6.2. Field Experiments

For further validating the practical working performance of the proposed spray boom control system, field experiments are carried out at Grain Industrial Park, Xinghua City, Jiangsu Province. The active boom control system is assembled on a tractor-mounted boom sprayer with a 24 m wide boom (3W-1200, Zhongnong Fengmao Plant Protection Machinery Co., Ltd., Beijing, China). The boom is suspended on the sprayer by a trapezoidal suspension.

The designed prototype is shown in Figure 5. To decrease crop damage, the tests are carried out in a slippery grass field 80 m long and 50 m wide. Narrow leaf *Setaria viridis* (L.) dominates in the field. There are from 8 to 15° slopes and blind pits close to 1 m. The fields are undulating, and the unevenness of the fields is 30–70 cm. When the sprayer operates in the field, the shaking of the pesticide tank (S_t) and the reaction force of the spray nozzle affect the vibration (S_v) of the boom chassis and boom. In this study, these effects are treated as external disturbances and non-modeling dynamics. The field experiments are also conducted with 3 replications. Before tests, the boom is lifted to a set height and kept parallel to the ground by manual adjustment. The control algorithm is started after the set values of the controlled parameters are set. Then, the posture of the boom is automatically adjusted when the sprayer travels about 4.8 km/h. In active balance mode, the inclination angles of the suspension chassis and boom are collected by two inclination sensors, which are assembled on the center boom and support chassis, respectively. The angle difference (θ_a) between the boom and the chassis is computed and used as the input variable. The corresponding output variable is the displacement (L_b) of the piston of the active balance cylinders. In the ground-following mode, under the action of the boom controller, the relative angle between the boom and the chassis is close to a small value, and the center boom keeps parallel to the crop canopy. Moreover, the boom end heights (H_b) are acquired by the ultrasonic sensors, which are installed on the two boom ends. By adjusting the displacement (output variable, L_i) of the piston of the imitation cylinders, the heights of the left-side boom arm and right-side boom arm maintain (input variables) close to the setting height under the action of the boom controller. By above the adjusting, the whole boom keeps the parallel to the crop canopy on ground. The definition of variables and factors for three different control modes is shown in Table 1.

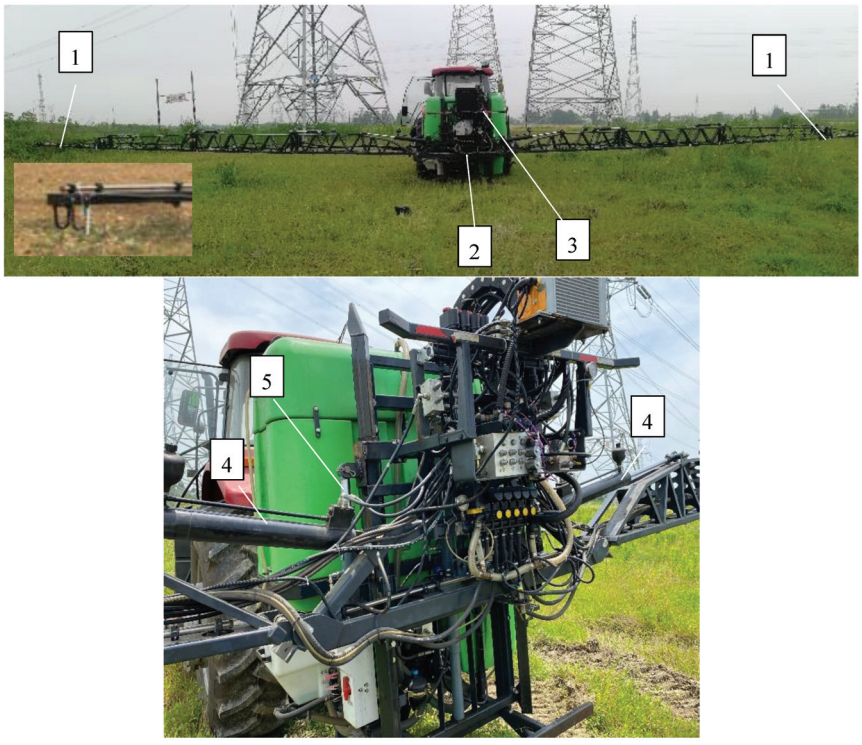


Figure 5. Tractor and boom suspension. 1—ultrasonic sensor; 2—inclination sensor of boom; 3—inclination sensor of chassis; 4—imitation cylinders; 5—active balance cylinder.

Table 1. Definition of variables and factors for three different control modes.

Control Mode	Input Variables	Output Variables	Factors
Active balance	θ_a	L_b	S_t, S_v
Ground following	H_b	L_i	S_t, S_v
Active balance and ground following	θ_a, H_b	L_b, L_i	S_t, S_v

3. Development of a Mathematical Model

It is vital to develop the dynamic model of the control object to facilitate performance analysis and accomplish control. The development of a mathematical model for boom chassis can be seen in Appendix A.

3.1. Analysis and Optimization of Boom-Passive Suspension Characteristics

Boom suspension is one of the most important parts in passive suspension, and the boom keeps the whole itself stable by relying on its own gravity. The advantage of this is that it hardly needs an external power source. The relationship between the structural parameters and damping of the passive suspension and the vibration characteristics of the boom is conducive to understanding of the vibration reduction characteristics of the boom under the passive suspension state. Meanwhile, this facilitates the analyze and optimization of the low-frequency following performance and high-frequency vibration isolation performance of the passive suspension. By analyzing and optimizing the low-frequency following performance and high-frequency vibration isolation performance of the passive suspension, the vibration damping performance in the passive state can be

further improved. Thus, the passive suspension is modified into an active suspension, and the energy consumption of the active suspension system can be reduced. Therefore, it is necessary to develop a dynamic model of passive boom suspension, and the performance of passive boom suspension can be optimized by analyzing the dynamic model.

3.1.1. Dynamic Model of an Active Balance System

The following assumptions are made for the passive suspension model:

(1) The boom suspension is a truss structure. The lubrication of the joints is good, and the friction can be ignored.
(2) The connections among the trusses are rigid, which is regarded as a rigid body, and the elastic deformations of all the trusses, which are in the vertical direction, are ignored.
(3) Neglecting the mass of the connecting parts of the truss.
(4) The inclination angle of the vehicle body is equivalent to the ground angle, and the error between the two angles is ignored.

The mathematical models of passive and active boom suspension are developed by analyzing the geometry and forces.

The force analysis of the passive suspension is shown in Figure 6. The boom is hinged to the boom suspension by twin link rods (AB and DC) symmetrically. The equivalent center of gravity of the boom is G, which is on the same vertical plane as the symmetry center of the boom. Moreover, twin active balance cylinders are mounted between the boom suspension and lift frame of the boom by hinge joints. Meanwhile, it is assumed that the vehicle body and boom rotate around the same O point on the ground when an incline from the ground happens during the process of traveling. Establish the coordinate system as shown in Figure 6. The origin of coordinate O is the projection point of the center of symmetry of the boom in the horizontal plane.

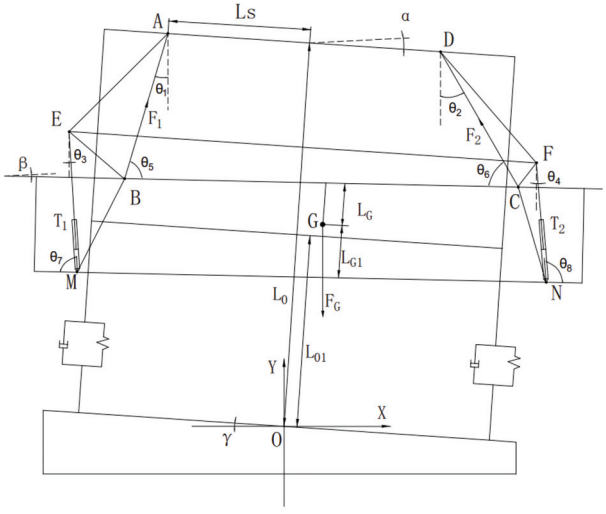


Figure 6. Boom suspension geometry.

The development of the dynamic model of an active balance system is as shown in Appendix A.

3.1.2. Dynamic Model of Ground Following Suspension

The geometric model of boom suspension is shown in Figure 7. The center of gravity and rotation center of the boom are O and H points, respectively. The cartesian coordinate system is established with point O as the origin of the coordinate, and the inclination angle

between the boom and horizontal plane is expressed by ϕ. On the one hand, the center boom is hinged to the right-arm boom by hinge joint R; on the other hand, two ends of the imitating hydraulic cylinder (RT) are hinged to the center boom and right-arm boom at R and T points, respectively. As the length of RT changes, the distance between the ground and the end of the spray boom varies nonlinearly, and the relationship between this distance and the variation in RT can be acquired by simulation.

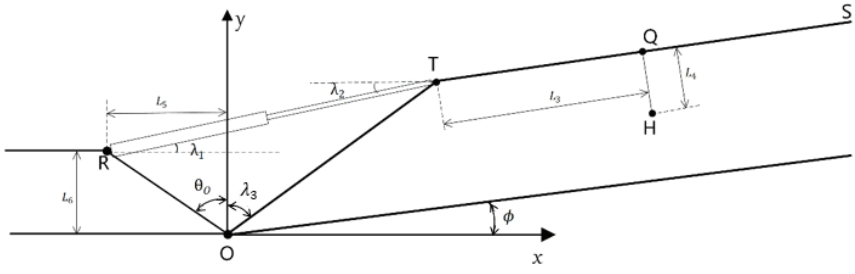


Figure 7. Ground following geometric model.

The dynamic equation of the ground-following mechanism can be obtained through the force analysis of the spray boom.

$$M_s\ddot{y} = -T_3 sin\lambda_1 + T_y - M_s g \tag{2}$$

$$M_s\ddot{x} = -T_3 cos\lambda_1 + T_x \tag{3}$$

$$I_s\ddot{\phi} = T_3 cos\lambda_1 L_y + T_3 sin\lambda_1 L_x + T_x y - T_y x \tag{4}$$

where

$$L_y = L_4 cos\phi - L_3 sin\phi \tag{5}$$

$$L_x = L_3 cos\phi + L_4 sin\phi \tag{6}$$

where L_3 is the distance between T and Q points; m; L_4 denotes the distance between Q and H points; m; ϕ is angle between the spray rod and the x axis; rad; T_3 is force of imitating hydraulic cylinder, N; T_x and T_y are the horizontal and vertical components of the force exerted on the spray rod at point B, respectively; M_s denotes the mass of the singleside boom; kg.

According to the expression of the center of mass coordinate H(x, y) of the boom, the Equations (7) and (8) can be derived from the projection relation in ΔROT.

$$\lambda_3 = -\theta_0 + arccos\frac{{L_{RO}}^2 + {L_{OT}}^2 - \left(L_{\overline{RT}} + x_L\right)^2}{2L_{RO}L_{OT}} \tag{7}$$

$$\lambda_1 = \theta_0 + arcsin\frac{{L_{OT}}^2 - {L_{RO}}^2 - \left(L_{\overline{RT}} + x_L\right)^2}{2L_{RO}\left(L_{\overline{RT}} + x_L\right)} \tag{8}$$

where λ_1 is the angle between imitating hydraulic cylinder and horizontal direction, rad; λ_3 angle between link rod BC and horizontal direction, rad; θ_0 angle between link rod RO and vertical direction, rad; x_L piston displacement of imitating hydraulic cylinder, m; $L_{\overline{RT}}$ initial length of cylinder barrel when $\phi = 0$, m; L_5 and L_6 are the horizontal and vertical projections of the distance between R and O, respectively, m. The relationship between x_L and the inclination angle ϕ of the boom as well as the force of the imitating cylinder can be obtained by Equations (2)–(8).

In the same way, the overall transfer function for the transducer, the valve, and the piston combination of the active imitating suspension can be rewritten as

$$G(s) = \frac{\Delta L_i}{\phi} = \frac{K}{s(Ts+1)} = \frac{B}{s^2 + As} \tag{9}$$

where ΔL_i denotes the piston displacement increment of the imitating hydraulic cylinder.

For the active balance system, $k_1 = \frac{K_q}{A_p} = \frac{1.275\times10^{-2}\ \mathrm{m^2{\cdot}s^{-1}}}{2.37\times10^{-3}\ m^2} = 5.38$ Hz, $k_2 = \frac{1}{\omega_h^2} = \frac{1}{15.39^2} = 4.22\times10^{-3}$; $k_3 = 2\times\frac{0.3}{15.39} = 0.039$. For the ground following control system, $k_1 = \frac{K_q}{A_p} = \frac{1.275\times10^{-2}\ \mathrm{m^2{\cdot}s^{-1}}}{6\times10^{-3}\ \mathrm{m^2}} = 2.125$ Hz, $k_2 = \frac{1}{\omega_h^2} = \frac{1}{6.32^2} = 0.025$, $k_3 = \frac{2\zeta_h}{\omega_h} = \frac{2\times0.2}{6.32} = 0.063$. The detailed expressions of Equations (A27) and (9) can be obtained by substituting these parameters. The linear relationship is obvious between ϕ and the distance from the boom to the ground. The expression of ΔL_i and the height (distance) from boom to ground can be obtained after the relationship between ΔL_i and ϕ is calculated by Equation (9). During the operating processing of the control systems, the heights from the left and right boom arms to the ground are measured in real time by the ultrasonic sensors 1 and 13 (as shown in Figure 1c) assembled on the left and right arms, respectively.

3.1.3. Control Algorithm

The main aim of active boom suspension is to maintain a desired spraying height from the crop canopy or ground in challenging terrain. Feedforward control is the control system that reacts to the variation in disturbance or set value according to the compensation principle; its characteristic is realizing the control action according to disturbance quantity so as to compensate for the influence of the disturbance on the controlled variable when the disturbance occurs, and the controlled variable has not changed before. Compared to feedback control, feedforward control can be controlled more timely and is hardly affected by system lag. Considering the hydraulic transfer delay of hydraulic oil in the lines and the response lag of actuators such as valves and hydraulic rams, feedforward proportion integration differentiation (FPID) control (Figure 8) is used to attempt to improve the tracking performance of the control system.

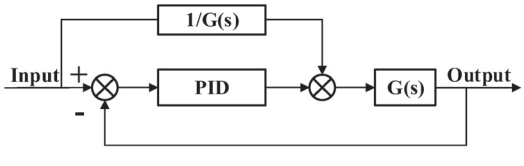


Figure 8. Block diagram of the control system.

According to the transfer functions G(s) of controlled objects in Equations (A27) or (9), the input-output expression for feedforward controllers can be derived as follows:

$$U_f(t) = \frac{A}{B}\frac{dx(t)}{dt} + \frac{1}{B}\frac{d^2x(t)}{dt^2} \tag{10}$$

where $x(t)$ denotes the input of the control system.

Discretize Equation (10), and the output of the feedforward link is rewritten as

$$U_f(k) = \alpha[x(k) - x(k-1)] + \sigma[x(k) - 2*x(k-1) + x(k-2)] \tag{11}$$

where $\alpha = A/(B*\mathrm{T_c})$, $\sigma = 1/\left(B*\mathrm{T_c^2}\right)$, $\mathrm{T_c}$ presents the control cycle. In this study, $\mathrm{T_c}$ was set at 50 ms.

The output of the PID controller based on feedforward compensation (FPID) is the sum of the traditional PID and feedforward compensation links.

$$U(k) = U_f(k) + U_p(k) \tag{12}$$

where $U_p(k)$ presents the output of PID. The PID parameters k_p, k_i, and k_d can be determined by trial and error.

During laboratory and field tests, the transfer functions $G(s)$ of controlled objects in Equations (A27) and (9) are integrated into the control algorithm (Figure 8), which is used to control the active balance control system and ground following control system, respectively.

4. Results and Discussion

4.1. Simulation and Experiment

To validate the effectiveness of the developed boom posture adjustment and control system, the simulation is performed by MATLAB software, and field experiments are carried out in the grain industrial park of Xinghua City, Jiangsu Province.

4.1.1. Expected Response Characteristics of a Passive Suspension

Damping and mechanical structure parameters are the main characteristic parameters of the passive suspension. In order to better analyze the influence of structural parameters and damping on the dynamic response characteristics, the expected characteristics of the passive suspension should be determined first.

The desired performance of passive suspension is isolating the high-frequency component and following the low-frequency component to maintain boom follow the undulating and sloping ground. The threshold value of frequency can be determined according to Ref [9].

$$\omega_{max} = 2\pi v / \lambda_{min} \tag{13}$$

where λ_{min} is wavelength of terrain, m; v denotes sprayer speed, $\mathrm{m \cdot s^{-1}}$; ω_{max} maximum tracking frequency of boom to terrain fluctuation, $\mathrm{rad \cdot s^{-1}}$.

Therefore, the passive suspension should transmit frequencies less than ω_{max} to the spray boom and attenuate frequencies greater than ω_{max}. Meanwhile, the time domain response amplitude of the boom suspension should decay to 10% of the initial value within 6 s as the boom is released freely from an inclined position [19].

4.1.2. Step Response Analysis

During the working process of the sprayer, if one side of the tire travels across the ridge of the bumpy field, the tractor chassis will tilt, which is equivalent to exerting a step excitation on the passive suspension. The simulation model of boom suspension (Figure 9) is established based on the mathematical model of trapezoidal passive suspension developed in the previous section. The input of the simulation model is a step signal. The step input is provided by moving a box with a certain height under the ultrasonic sensor 1 or 13 (Figure 1). The diameter of the valve orifice is an important parameter that influences the response speed and oscillations of the suspension and boom suspension. Three different valve's orifices of valve 13 (Figure 2) are set manually to understand the response characteristic of boom suspension at various working parameters. The corresponding parameters of boom suspension are shown in Table 2. The step response is presented in Figure 10 for three different diameters of throttling orifice holes.

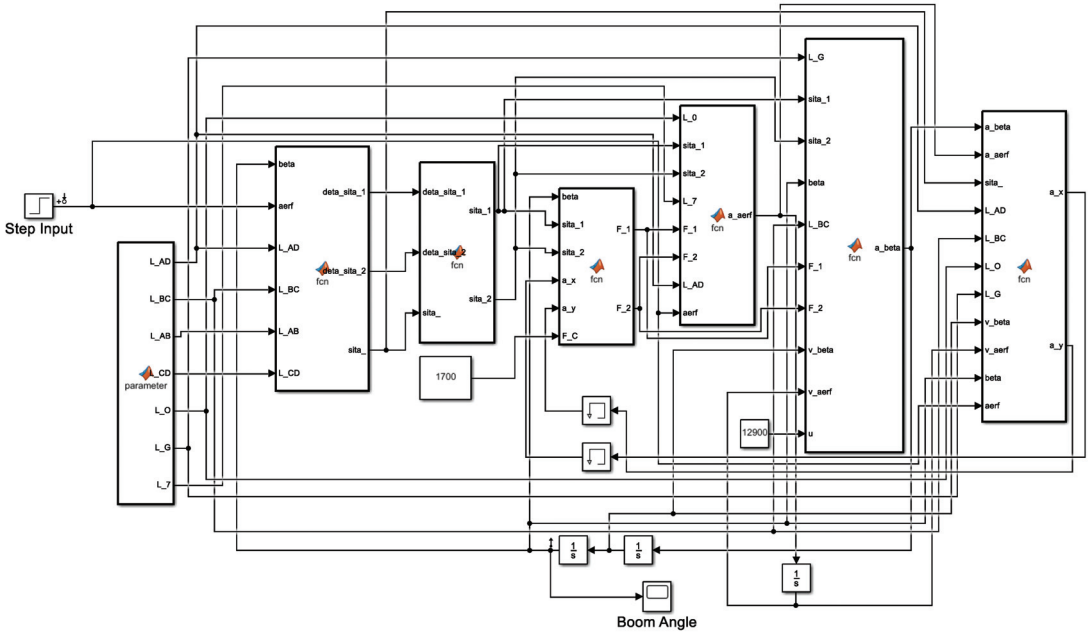


Figure 9. Simulation model of passive boom suspension.

Table 2. Parameters of passive boom suspension.

Sign	Parameter Signature	Parameter Value	Unit
L_{G1}	Length of boom gravity to MN	0.198	m
L_G	Length of boom gravity to BC	0.200	m
$\overline{L}_{EM}$	Initial length of left active balance hydraulic cylinder	0.520	m
$\overline{L}_{FN}$	Initial length of right active balance hydraulic cylinder	0.520	m
L_{MN}	Length of MN	1.420	m
L_{EF}	Length of EF	1.344	m
I_B	Moment of inertia of boom about mass center	1.488×10^4	kg·m^2
M	Mass of boom	535	kg
$\overline{\theta}$	Initial angle between AB/DC and vertical direction	1.05–0.52	rad
L_0	Length shown in Figure 4	2.00	m
L_{AB}	Length of AB	0.37	m
L_{CD}	Length of CD	0.37	m
L_{AD}	Length of AD	0.48	m
L_{BC}	Length of BC	0.80	m

Figure 10 implies that the diameter of the orifice has a great influence on the peak value of the step response and the number of oscillations. As the orifice diameter drops, the system damping climbs. Thus, the passive suspension of the boom can reach a stable state more quickly, and the number of oscillations and the maximum overshoot will be reduced. However, if the throttle orifice is set too small, the time to reach the peak value increases, resulting in an increase in the response time of the suspension and a slow response speed. In the above analysis, the orifice diameter of 5 mm is determined for the design of the hydraulic system.

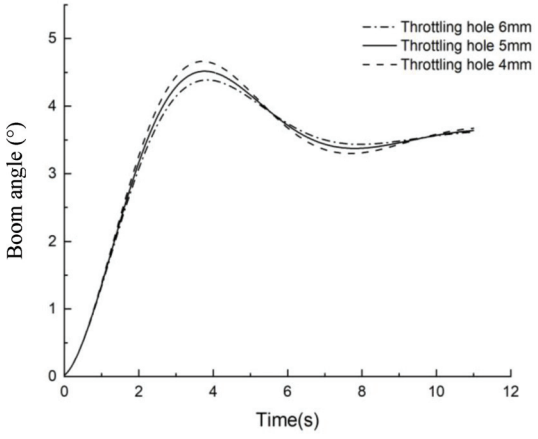


Figure 10. Step response characteristic curve of boom suspension.

4.1.3. Frequency Domain Response Analysis of Passive Suspension

The second-order model of passive boom suspension can be obtained from the dynamic model of passive suspension when the inclination angle β is small.

$$C\ddot{\beta} = A\ddot{a} + \mu_B\left(\dot{\alpha} - \dot{\beta}\right) + B\alpha - D\beta \tag{14}$$

Considering weakly damped structure and ignoring the mechanical friction damping of boom suspension, Equation (14) can be simplified as follows:

$$C\ddot{\beta} = A\ddot{a} + B\alpha - D\beta \tag{15}$$

where μ_B is boom suspension rotational damping coefficient, N m s/rad;

$$\begin{gathered} A = M\left[L_0 + \frac{L_{BC}}{2\tan\overline{\theta}} - L_{AB}\cos\overline{\theta}\right]\left[L_G + \frac{L_{BC}}{2\tan\overline{\theta}}\right] \\ B = Mg\left[\frac{(L_{BC}/2)\left[(L_{BC}/2) - L_{AB}\sin^3\overline{\theta}\right]}{L_{AB}\cos\overline{\theta}\sin^2\overline{\theta}} - \frac{L_{BC}}{2\tan\overline{\theta}}\right] \\ C = I_B + M[L_G + \frac{L_{BC}}{2\tan\overline{\theta}}]^2 \\ D = Mg\left[\frac{(L_{BC}/2)\left[(L_{BC}/2) - L_{AB}\sin^3\overline{\theta}\right]}{L_{AB}\cos\overline{\theta}\sin^2\overline{\theta}} + L_G\right] \end{gathered}$$

By applying the Laplace transformation to Equation (15), the transfer function of the boom inclination β and the boom suspension inclination α can be acquired.

$$\frac{\beta}{\alpha} = H(s) = \frac{As^2 + B}{Cs^2 + D} \tag{16}$$

From Equation (16), it can be found that passive suspension belongs to the second-order system, and the characteristic equation of the system is expressed as

$$Cs^2 + D = 0 \tag{17}$$

The natural frequency of the system, $\omega_n\left(\text{rad}\cdot\text{s}^{-1}\right)$ is described as

$$\omega_n = \left(\frac{D}{C}\right)^{0.5} \tag{18}$$

The main function of passive suspension is to attenuate the excitation from the tractor chassis and road roughness. Equation (16) implies the response characteristics of the passive suspension can be adjusted by the length of twin link rods. The natural frequencies of boom suspension are 0.74 rad $\cdot$ s^{-1}, 1.086 rad $\cdot$ s^{-1} and 1.57 rad $\cdot$ s^{-1} when the lengths of twin link rods are set as 0.38 m, 0.48 m, and 0.58 m, respectively. The corresponding frequency response plot (Figure 11) means the natural frequency increases as twin-link rod length increases, and the amplitude ratio increases slowly and then decreases rapidly to almost the same value. For different initial lengths of link rod, the amplitude ratio of the passive suspension of the spray rod differs greatly at low frequencies. When the frequency exceeds 1.086 rad$\cdot$s^{-1}, the amplitude ratio decays quickly, which shows the characteristics of high-frequency isolation of the passive suspension.

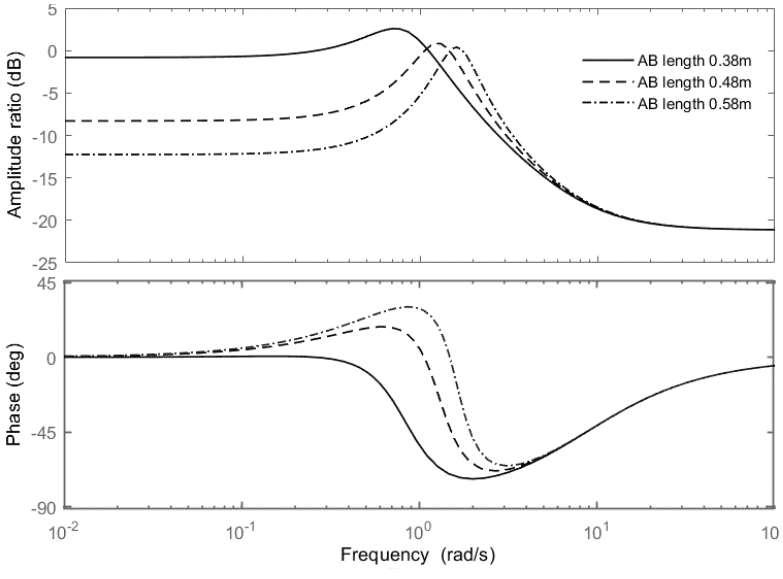


Figure 11. Frequency response characteristics of passive suspension with different AB lengths.

From the phase-frequency response plot of the passive suspension at the natural frequency, it can be seen that the phase becomes more laggy with the increase in the length of the connecting rod. The length of 0.48 m is determined for twin link rods in mechanical structure design.

4.1.4. Ground-Following Dynamics Simulation

The simulation for the mathematical model of the ground following the development in the previous section is performed in Matlab/Simulink (Figure 12). The rungkutta numerical integration algorithm is used to solve the differential equations. The correspond−ing simulation results can be seen in Figure 13. Finally, the relationship between cylinder displacement x_p and the change in nozzle end height H can be obtained through angle transformation (Figure 14).

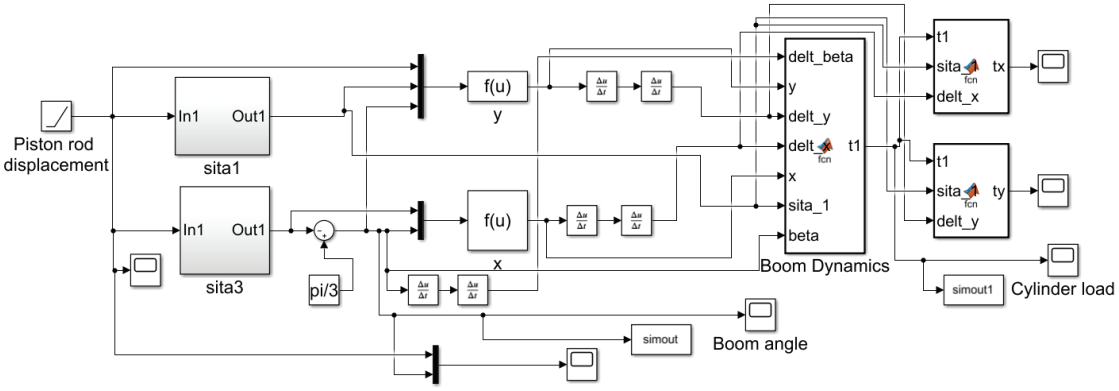


Figure 12. Dynamic simulation model for ground following of the boom suspension.

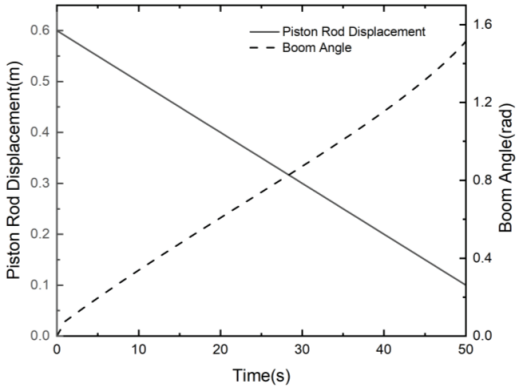


Figure 13. Relationship between cylinder displacement x_p and β.

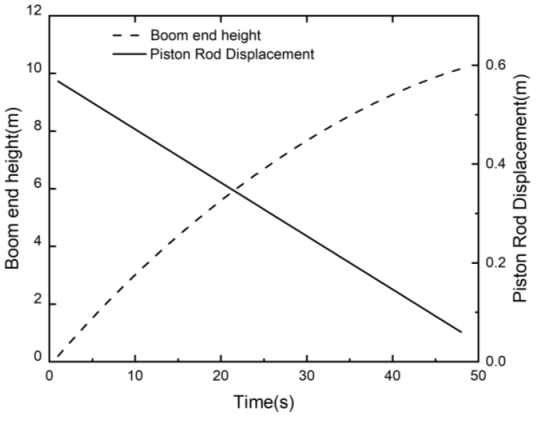


Figure 14. Relationship between cylinder piston displacement and nozzle end height.

4.1.5. Evaluation of Response Times of the Developed Hydraulic System

The response time of the hydraulic system, which mainly depends on the performance of proportional valves and pipeline lengths between the proportional valves and cylinders, is an important index for evaluating the working efficiency. Considering different pipeline lengths, the reaction times of the ground following the hydraulic system and the active

balance hydraulic system are evaluated, respectively. The experiments are designed with 5 replications. First, by inputting various control currents, the opening of the proportional valve is altered, ranging from 10% to 70% (step size 10%) under working pressure of 7 Mpa. Then, the time from the input signal change of the proportional valve to the output signal change of the hydraulic cylinder is recorded by the controller, and the average values of 5 replications are used as the response time. Experimental results indicate that the response time of the ground-following hydraulic system ranges from 0.46 s to 0.62 s, while that of the active balance hydraulic system ranges from 0.39 s to 0.52 s. This implies that the designed hydraulic system has fast response characteristics.

The approximately linear relationship between piston rod displacement of the hydraulic cylinder and the boom end height is found as shown in Figure 14. Within the cylinder stroke range of 0 to 60 cm, the boom end height varied from 0 to 10.6 m. During this process, the maximum force of the imitating cylinder is 4.12×10^4 N (Figure 15). The inversely proportional relationship can be seen between the maximum force of the imitating cylinder and the inclination angle β.

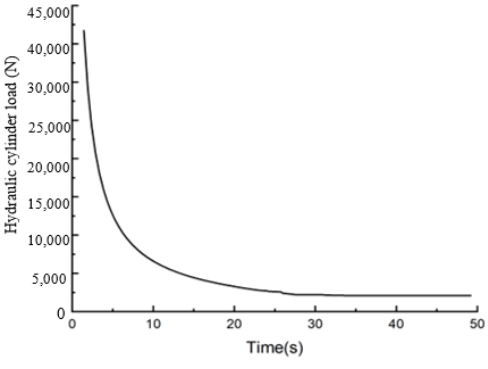


Figure 15. Force variation in an imitating cylinder.

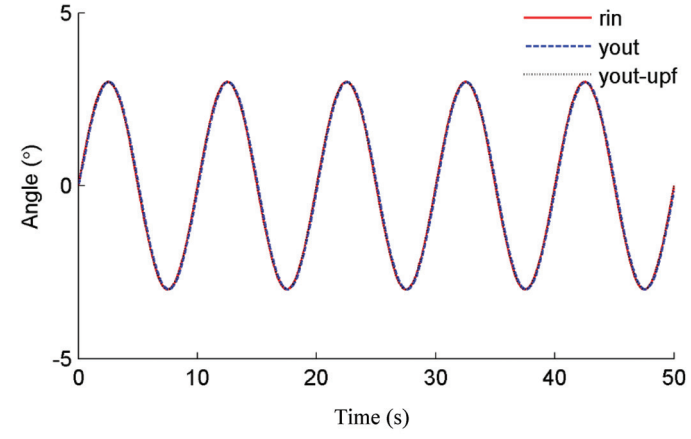
4.2. Experiments and Results

4.2.1. Testing of the Active Balance Control System

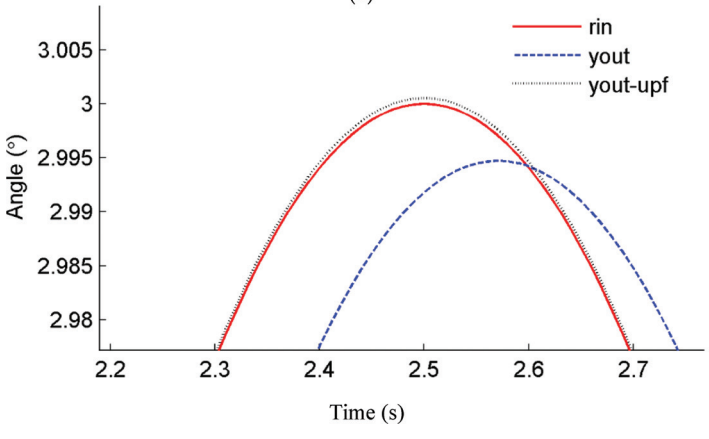
The boom suspension was fixed to a six-degree-of-freedom hydraulic motion platform, and the proposed boom control system was set to active balance mode. The experiments were carried out in the laboratory of Jiangsu University. The control parameters of the FPID and PID are $k_p = 8.0$, $k_i = 0.5$, $k_d = 1.5$, determined by trial and error, which denote proportional gain, integrating gain, and differential gain, respectively. The same sinusoidal stimulation $x(t) = 3\sin(0.2\pi t)$ is output by a six-degree-of-freedom hydraulic motion platform. The PID and FPID control algorithms are operated by the active balance control system. The system outputs with the PID and FPID controllers are shown in Figure 16a, and the corresponding tracking errors are presented in Figure 16c. The performance indexes during the last ten control cycles are presented in Table 3. M_a, μ_a, and σ denote the maximum, average, and standard deviation of the absolute errors, respectively.

Table 3. Performance indicators under the action of two controllers.

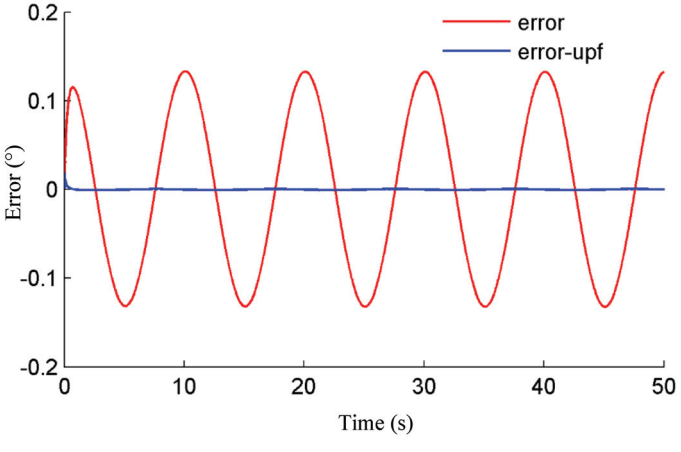
Indexes	M_a/°	μ_a°	σ/°
PID	0.12	0.043	0.039
FPID	0.03	0.012	0.009



(**a**)



(**b**)



(**c**)

Figure 16. Comparison of two controllers for an active boom control system. (**a**) Tracking performance for two controllers; rin: input signal; yout: PID; yout-upf: FPID; (**b**) Local amplification of (**a**); (**c**) Tracking error for the two controllers.

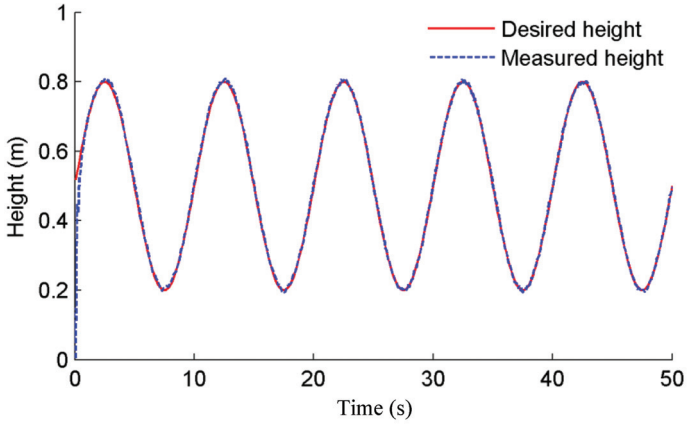
Figure 16b shows that the system output (yout) based on the PID controller lags the input signal (rin) while the system output(yout-upf) based on the FPID control algorithm obtains good tracking performance. The maximum steady-state training error of PID is 0.12°, while the maximum steady-state training error of FPID is 0.03°. As a kind of feedback control, PID adjusts the controlled quantity according to deviation. Under the action of interference, the controlled quantity first deviates from the given value, and then the regulator exerts control action to counteract the effect of interference. Thus, the system output always lags the given value. Moreover, according to the basic principles of feedback control, it is inevitable that there will be steady-state position tracking errors. The feedforward control is an open-loop link that acts according to the compensation principle of the disturbance or the change of the given value. Proper use of the feedforward control system can eliminate the disturbance of the controlled variable in the bud, so that the controlled variable will not be deviated by the disturbance effect or the change of the given value. Compared with the feedback control, it can be controlled in a more timely way and is not affected by the system lag.

4.2.2. Testing of the Ground Following the Control System

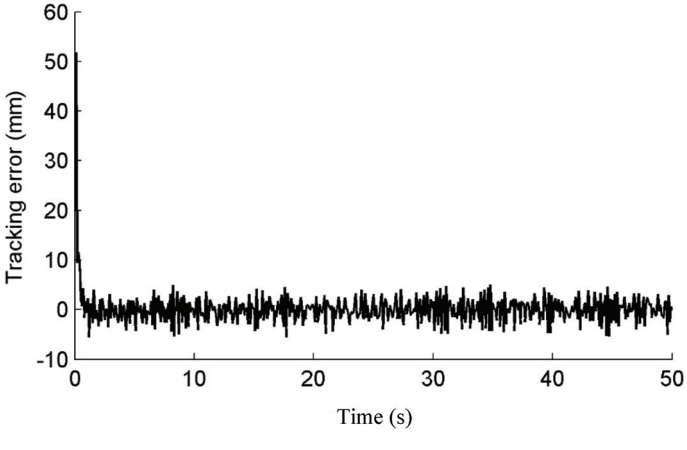
To achieve more consistent droplet distribution and reduce spray drift, it is expected to keep the spray boom parallel to the crop canopy or ground and maintain a certain distance from the spray nozzles to the crop canopy or ground.

To validate the performance of boom automatic following with crop canopy or ground in sloping and undulating terrains, tests for the ground following control system are conducted. The proposed boom control system was set to ground-following mode. A six-degree-of-freedom hydraulic motion platform is adopted to simulate sloping and undulating terrains, and the performance of the FPID controller is evaluated. The control parameters of the FPID controller are $k_p = 8.0$, $k_i = 0.5, k_d = 0.5$. The sine signal $h(t) = [0.3\sin(0.2\pi t) + 0.5]\ m$ is the input of the ground following the control system. During the adjustment processing of the control systems, the heights from left and right boom arms to ground are recorded by the ultrasonic sensors 1 and 13 (as shown in Figure 1c) assembled on the left and right arms, respectively. The imitating performance of the control system can be seen in Figure 17a. The tracking error can be found in Figure 17b. The maximum absolute error and the standard deviation (SD) are 0.063 m and 8.7 mm, respectively. As can be seen from Figure 17, the measured boom end height is close to the desired height, which implies that the ground following control system has superior following performance and can keep the spray boom parallel to the crop canopy or ground under complex and varied terrain conditions.

In the ground-following control system, the rolling motion of one arm will cause the gravity center change of the entire spray boom and then affect the other arm. It is neces− sary to consider the coupling action between the two boom arms. To solve this problem, we adopted two shutoff valves in the design of the hydraulic system. Each solenoid valve can be controlled by the controller individually. For an active balance control system, the two shutoff valves are opened, and hydraulic oil can pass through the valves. For the ground-following control system, when the boom arms are adjusted, the two shutoff valves are closed, and hydraulic oil is locked in the hydraulic pipes. By the above analysis, good performance of the control system can be seen, and the designed system is effective for less challenging terrain.



(**a**)



(**b**)

Figure 17. Tracking performance for the ground following the control system. (**a**) Desired height and measured height; (**b**) imitating error.

4.2.3. Testing of the Ground Following the Control System for Challenging Terrain

Due to the inconsistent soil firmness and bump pit in the field, the single-side tires of the tractor trap into the ground during the spraying of the sprayers, which tilt the boom chassis and cause the boom end to touch the ground. In this unexpected circumstance, the boom wing must be raised as rapidly as possible to avoid damaging the boom. To evaluate the performance of the control system in challenging terrain, a step-response test is carried out. The control parameters of the FPID controller are $k_p = 8.0$, $k_i = 0.5$, $k_d = 0.5$. Firstly, the sprayer boom is adjusted to a horizontal position and is 0.8 m above the ground. Then the control system started. A step signal is generated by rapidly moving the carton (0.5 m high) to the bottom of the ultrasonic sensor assembled on the right arm. During the adjustment processing of the control systems, the heights from left and right boom arms to ground are recorded by the ultrasonic sensors 1 and 13 (as shown in Figure 1c) assembled on the left and right arms, respectively. The height difference between the left boom arm and the right boom arm is calculated. The fluctuation of the height difference between the left boom arm and the right boom arm is shown in Figure 18. Two peaks can be observed from Figure 18. At the beginning, the heights between left and right boom arms

and ground are both 0.8 m, and the height difference between left and right boom arms is close to 0. At about 5 s, a step signal is generated by rapidly moving the carton (0.5 m in height) to the bottom of the ultrasonic sensor assembled on the right arm. At this moment, the ultrasonic sensor assembled on the right boom arm senses the distance between boom and carton, and this distance is 0.3 m, which is less than the pre-setting height of 0.8 m. Meanwhile, the height between the left boom arm and the ground is 0.8 m. Thus, the height difference between the left and right boom arms is equal to 0.5 m. To regain the pre-setting height between the spray boom and the carton, the right boom arm gradually lifts under the action of the control system. During this process, the height between the left boom arm and the ground still maintains 0.8 m, which results in the height difference between the left and right boom arms gradually decreasing. The height difference between the left and right boom arms is close to 0 again when the distance between the boom and carton reaches 0.8 m.

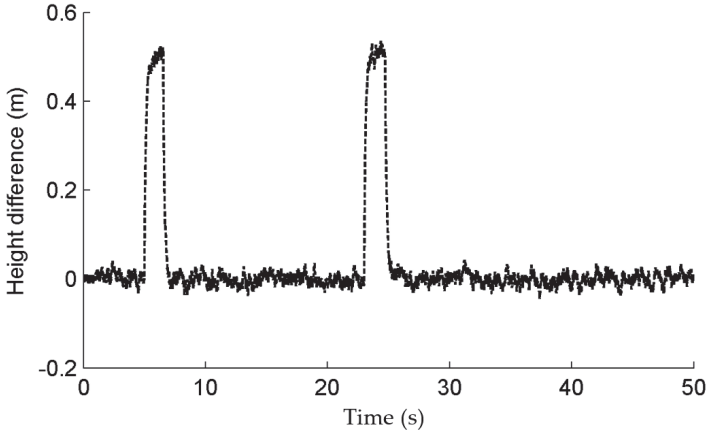


Figure 18. Height difference between left and right boom arms.

When rapidly moving the carton away from the bottom of the ultrasonic sensor assembled on the right arm, the ultrasonic sensor assembled on the right boom arm end senses the distance between boom and ground, and this distance is 1.3 m. The corresponding height difference between the left and right boom arms is -0.5 m. The right boom arm gradually falls under the action of the control system, which results in the height difference between the left and right boom arms gradually decreasing. The height difference between the left and right boom arms is close to 0 again when the distance between boom and ground reaches 0.8 m. Compared with the lifting process, the dropping process of the boom is faster. Considering the main aim of this test is to evaluate the response performance of the proposed control system under challenging conditions, the data on the falling process of the spray boom is not presented in this manuscript.

By the above analysis, the peak means the height difference between the left and right boom arms for the moment of moving the carton to the bottom of the ultrasonic sensor assembled on the right arm. To adequately validate the effectiveness of the control system, several repeated experiments are carried out. Figure 18 presents the experimental data for two replications. For the peak at about 23 s (second replication), which is the same as that at about 5 s.

From Figure 18, it can be seen that the difference between the left and right boom arms drops to less than 10% of the initial distance within 1.87 s by the excitation of the step signal at about 5 s and 23 s. This indicates that the proposed control system can achieve precise control in challenging terrain, which ensures the boom end hardly touches the ground or crop canopy under challenging circumstances such as the deflection of the vehicle tires and sinkage of wheels into the ground.

4.2.4. Field Experiment

According to the description of the field experiment in Section 2.5, the field experiments are conducted aiming to evaluate the performance of the boom-controlled system in the hybrid mode of active balance and ground following. The rolling angle variation in the boom and chassis is shown in Figure 19. The statistical analysis of boom inclination angle and chassis inclination angle is given in Table 4.

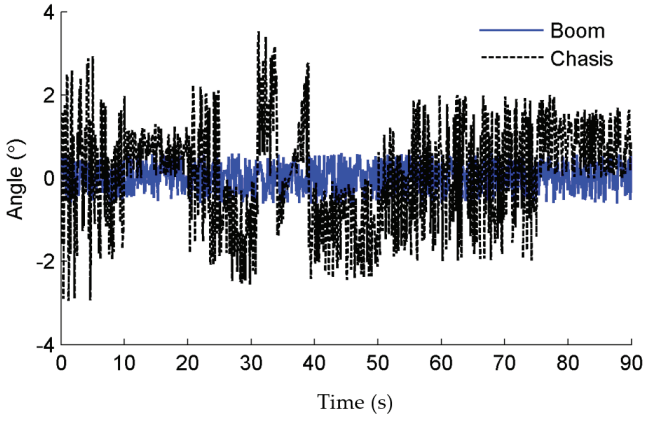


Figure 19. Rolling angle variation in boom and chassis in a field experiment.

Table 4. Statistical analysis of boom inclination angle and chassis inclination angle.

Indexes	M_a/°	μ_a°	σ/°
Chasis	3.896	0.984	0.826
Boom	0.453	0.130	0.117

As can be seen from Figure 19, the chassis waggles violently from the excitation from the soil deformation and bumpy road when the sprayer travels across the slippery and undulating soil. The maximum rolling angle of the chassis is 3.896°. Under the regulating action of the proposed control system, the inclination angle of the boom maintains a relatively small range, and the maximum angle is 0.453°. It is proven that the designed control system can maintain the boom parallel to the crop canopy or ground while isolating the high-frequency vibration from the chassis. The reason is that the forward feedback control can directly exert the correction effect according to the disturbance quantity when the disturbance occurs. Thus, the VPID effectively eliminates the external disturbance from chassis vibration before the controlled variable has changed. In addition, the precise system model has the premise of eliminating the deviation caused by disturbance, which implies that the developed mathematical models are appropriate and accurate.

A boom control system was developed for traditional pendulum active suspension [15], and field tests were conducted by assembling the developed control system on a 12 m wide boom. Results showed that the maximum roll angle of the boom and the maximum roll angle of the chassis were 0.841° and 4.041°, respectively. By comparison, it can be found that the designed control system in this study, whose obtained performance is slightly higher than that of Ref. [15]. The reason is probably that accurate mathematical models are obtained, which meet the requirements of accurate mathematical models needed for feedforward control.

5. Conclusions

A high-efficiency boom posture adjustment and control system was designed based on the FPID algorithm, which was superior to traditional trapezoid active suspension.

The proposed active balance control system not only realized the automatic balance of boom in flat terrains and boom automatic following with crop canopy or ground in sloping and undulating terrains but also achieved the quick raise of boom wing in emergency circumstances such as unilateral tire of sprayer sinkage into the ground quickly and avoid damage because of touching between boom tip and target area (soil and crop canopy). Compared with available active boom control systems, the control system in this study has advantages such as rapid response characteristics, high operating efficiency, and safety.

To clarify the relationship among the structural parameters and damping of the passive suspension and the vibration characteristics of the boom and optimize the low-frequency following performance, the dynamic model of passive boom suspension was developed. The step response characteristic curve of boom suspension for three different orifice diameters and the frequency response characteristics of passive suspension with three different AB lengths were analyzed by simulations. Moreover, to improve the tracking performance of the control system, the FPID control algorithm was designed according to the dynamic model of the active boom system.

The designed control system was assembled on a 24 m boom with trapezoid suspension. The response characteristic of the active boom control system was tested by the step signal and the sinusoidal signal from a six-degree-of-freedom hydraulic motion platform. Firstly, the tracking performance of the active balance control system for the PID and FPID control algorithms was compared for a given 0.2 Hz sine signal. Then, for the ground-following control system, the response characteristics in challenging terrain and tracking performance in less challenging terrain were tested. Field experiment results indicated that the maximum rolling angle of the chassis was 3.896°, while the maximum inclination angle of the boom was 0453°. The results show that the designed control system can effectively adjust the boom motion in real time and meet the requirements of field operation. Meanwhile, the simulation results were validated by the field tests. In this study, the energy consumption of the proposed control system is not considered, and it is necessary to evaluate it in subsequent research.

The novelty of this study is that it provides an automatic control device with low-cost and high-response performance of spaying height for trapezoidal active suspension, aiming to maintain the desired spraying height from crop canopy or ground in challenging terrain. The main contributions are improving the response performance of boom control devices in challenging terrain and providing a solution to developing a low-cost boom altitude controller that matches the available sprayer. The proposed design methodology, including the hydraulic system and hardware circuit of the boom control system, the performance test of the boom suspension, and the determination of key parameters of the boom suspension and hydraulic system, provides the theoretical basis and practical experience for future low-cost and intelligent boom active control systems. It is of great significance to optimize and upgrade the high-performance active boom suspension in the future.

Author Contributions: Conceptualization, J.L. and Z.N.; methodology, D.G.; software, Z.N.; validation, J.L.; formal analysis, Y.C.; investigation, J.L.; resources, J.L.; data curation, Y.C.; writing—original draft preparation, D.G.; writing—review and editing, M.L.; visualization, M.L.; supervision, J.L.; project administration, J.L. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by the Key Research and Development Program of Zhenjiang city (NY2022008), and the Key Research and Development Program of Jiangsu Province (BE2018324).

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: Not applicable.

Conflicts of Interest: The authors declare no conflict of interest.

Appendix A

Appendix A.1. Mathematic Model of Valve-Controlled Hydraulic Cylinder

The flow equation of an electro-hydraulic proportional valve can be described as

$$q_L = K_q x_v - K_c p_L \tag{A1}$$

where q_L is load flow of electro-hydraulic proportional valve, $\left(\mathrm{m}^3 \cdot \mathrm{s}^{-1}\right)$; x_v spool displacement of proportional valve; m; K_c flow-pressure coefficient of proportional valve, $\mathrm{m}^5 \cdot \mathrm{N}^{-1} \cdot \mathrm{s}^{-1}$; K_q flow gain coefficient of proportional valve, $(\mathrm{m}^2 \cdot \mathrm{s}^{-1})$; p_L loading pressure, (Pa).

The flow continuity equation of a hydraulic cylinder can be expressed as

$$q_L = A_p \frac{dx_p}{dt} + C_{tm} p_L + \frac{V_t}{4\beta_e} \frac{dp_L}{dt} \tag{A2}$$

where A_p is hydraulic cylinder cross-sectional area, m^2; x_p denotes piston rod displacement of hydraulic cylinder, m; C_{tm} total leakage coefficient of hydraulic cylinder, $\mathrm{m}^5 \cdot \mathrm{N}^{-1} \cdot \mathrm{s}^{-1}$, $C_{tm} = C_{im} + \frac{1}{2} C_{em}$; ($C_{im}$ and C_{em} are internal and external leakage coefficients, respectively); V_t total volume of hydraulic cylinder, proportional valve chamber, and connecting pipes, m^3; β_e effective volumetric elastic modulus of hydraulic oil, Pa.

Ignoring the nonlinear static friction, Coulomb friction force, and oil mass, the torque balance equation of the hydraulic cylinder and load based on Newton's second law can be obtained.

$$A_p p_L = m_t \frac{d^2 x_p}{dt^2} + B_p \frac{dx_p}{dt} + G x_p + F_L \tag{A3}$$

where m_t is the total mass of piston and load, kg; B_p viscous damping coefficient, $\mathrm{N} \cdot \mathrm{s} \cdot \mathrm{m}^{-1}$; G load spring stiffness, $\mathrm{N} \cdot \mathrm{m}^{-1}$; F_L load force, N.

By taking the Laplace transform of Equations (A1)–(A3), the piston displacement of the hydraulic cylinder can be obtained:

$$x_p(s) = \frac{\frac{K_q}{A_p} x_v(s) - \frac{K_{ce}}{A_p^2}\left(1 + \frac{V_t}{4\beta_e K_{ce}} s\right) F_L(s)}{\frac{m_t V_t}{4\beta_e A_p^2} s^3 + \left(\frac{B_p V_t}{4\beta_e A_p^2} + \frac{m_t K_{ce}}{A_p^2}\right) s^2 + \left(1 + \frac{G V_t}{4\beta_e A_p^2} + \frac{B_p K_{ce}}{A_p^2}\right) s + \frac{G K_{ce}}{A_p^2}} \tag{A4}$$

where K_{ce} ($K_{ce} = K_c + C_{tm}$), is the total flow-pressure coefficient, $\mathrm{m}^5 \cdot \mathrm{N}^{-1} \cdot \mathrm{s}^{-1}$.

Considering the load characteristic of boom sprayers is inertial load, the effect of elastic load is negligible ($G = 0$). In addition, $\frac{B_p K_{ce}}{A_p^2} << 1$, Equation (A4) can be simplified as

$$x_p(s) = \frac{\frac{K_q}{A_p} x_v(s) - \frac{K_{ce}}{A_p^2}\left(1 + \frac{V_t}{4\beta_e K_{ce}}\right) F_L(s)}{s\left(\frac{s^2}{\omega_h^2} + \frac{2\zeta_h}{\omega_h} s + 1\right)} \tag{A5}$$

where ω_h is undamped hydraulic natural frequency, $\omega_h = \sqrt{\frac{4\beta_e A_p^2}{V_t m_t}}$, $\mathrm{rad} \cdot s^{-1}$; ζ_h is hydraulic damping ratio, $\zeta_h = \frac{K_{ce}}{A_p}\sqrt{\frac{\beta_e m_t}{V_t}} + \frac{B_p}{4A_p}\sqrt{\frac{V_t}{\beta_e m_t}}$ is dimensionless parameter.

Near the steady operating point of the system, $Q_L(s) = K_q X_v(s)$. The following transfer functions can be obtained based on Equation (A5).

$$\frac{x_p(s)}{Q_L(s)} = \frac{\frac{1}{A_p}}{s\left(\frac{s^2}{\omega_h^2} + \frac{2\zeta_h}{\omega_h} s + 1\right)} \tag{A6}$$

$$\frac{x_p(s)}{F_L(s)} = \frac{\frac{K_{ce}}{A_p^2}\left(1 + \frac{V_t}{4\beta_e K_{ce}}s\right)}{s\left(\frac{s^2}{\omega_h^2} + \frac{2\zeta_h}{\omega_h}s + 1\right)} \tag{A7}$$

The transfer function of a proportional valve can be described as

$$G_v(s) = \frac{Q(S)}{I(S)} = \frac{K_q}{\frac{s^2}{\omega_v^2} + \frac{2\zeta_v s}{\omega_v} + 1} \tag{A8}$$

where ω_v and ζ_v are proportional valve natural frequency and damping ratio, respectively.

The transfer functions of hydraulic cylinder speed to flow and external load can be obtained, respectively.

$$\frac{\dot{x}_p(s)}{Q(s)} = \frac{\frac{1}{A_p}}{\frac{s^2}{\omega_h^2} + \frac{2\zeta_h}{\omega_h}s + 1} \tag{A9}$$

$$\frac{\dot{x}_p(s)}{F_L(s)} = \frac{\frac{K_{ce}}{A_p^2}\left(1 + \frac{V_t}{4\beta_e K_{ce}}s\right)}{\frac{s^2}{\omega_h^2} + \frac{2\zeta_h}{\omega_h}s + 1} \tag{A10}$$

where $\frac{K_{ce}}{A_p^2}\left(1 + \frac{V_t}{4\beta_e K_{ce}}s\right) = K_T$.

Considering the angle sensor as a proportional element, the transfer function of the displacement sensor can be written as

$$K_f = \frac{U(s)}{\dot{x}_p(s)} \tag{A11}$$

where K_f is angle sensor gain, $\mathrm{V \cdot s \cdot m^{-1}}$.

The E-ME-L-01 type proportional amplifier is used with the electro-hydraulic proportional direction valve. The voltage-rated input value is $U_0 = 10(\mathrm{V})$ and the corresponding rated output current is $I_0 = 3(\mathrm{A})$, therefore, proportional amplifier gain K_a can be expressed as

$$K_a = \frac{I_0}{U_0} = \frac{3}{10} = 0.3\left(\mathrm{A \cdot V^{-1}}\right) \tag{A12}$$

Appendix A.2. Dynamic Model of an Active Balance System

According to the projection relationship, the coordinate (x, y) of the center of mass of boom G relative to O can be described as:

$$\begin{cases} x = L_0 sin\alpha + \frac{1}{2}L_{\mathrm{AD}} cos\alpha + L_{\mathrm{CD}} sin\theta_2 - \frac{1}{2}L_{\mathrm{BC}} cos\beta - L_{\mathrm{G}} sin\beta \\ x = L_0 sin\alpha - \frac{1}{2}L_{AD} cos\alpha - L_{\mathrm{AB}} sin\theta_1 + \frac{1}{2}L_{\mathrm{BC}} cos\beta - L_{\mathrm{G}} sin\beta \end{cases} \text{(lean to left)} \tag{A13}$$

$$\begin{cases} y = L_0 cos\alpha + \frac{1}{2}L_{\mathrm{AD}} sin\alpha - L_{\mathrm{AB}} cos\theta_1 - \frac{1}{2}L_{\mathrm{BC}} sin\beta - L_{\mathrm{G}} cos\beta \\ y = L_0 cos\alpha - \frac{1}{2}L_{\mathrm{AD}} sin\alpha - L_{\mathrm{CD}} cos\theta_2 + \frac{1}{2}L_{\mathrm{BC}} sin\beta - L_{\mathrm{G}} cos\beta \end{cases} \text{(lean to right)} \tag{A14}$$

where L_{AD}, L_{AB}, L_{BC} and L_{CD} are the AD, AB, BC and CD lengths, respectively, m; L_o vertical distance between the center of the link AD and the ground, m; α and β are inclination angles of suspension and boom relative to the horizontal plane, respectively, rad; L_G the vertical distance from the center mass of boom to BC, m.

In trapezoidal suspension ABCD, the initial Angle between the vertical direction and link rods AB and DC both are $\overline{\theta}$, thus $\overline{\theta}$ can be expressed as:

$$\overline{\theta} = arcsin\left(\frac{L_{\mathrm{BC}} - L_{\mathrm{AD}}}{L_{\mathrm{AB}} + L_{\mathrm{CD}}}\right) \tag{A15}$$

In the process of boom movement, given that the change range of α, β, θ_1 and θ_2 are relatively small, the coordinate (x, y) can be expressed as:

$$\begin{cases} x = \left(L_0 + \frac{cos\bar{\theta}}{2sin\bar{\theta}} L_{\mathrm{AD}}\right)\alpha - \left(L_{\mathrm{G}} + \frac{cos\bar{\theta}}{2sin\bar{\theta}} L_{\mathrm{BC}}\right)\beta \\ y = L_0 - L_{\mathrm{G}} - \frac{1}{2}\left(L_0\alpha^2 - L_{\mathrm{G}}\beta^2\right) + \frac{cos\bar{\theta}}{4sin\bar{\theta}}\left(L_{\mathrm{AD}}\alpha^2 - L_{\mathrm{BC}}\beta^2\right) \end{cases} \tag{A16}$$

$$\begin{cases} \Delta\theta_2 \approx \frac{1}{2L_{CD}}\left[\frac{L_{AD}\alpha^2}{2cos\bar{\theta}} - \frac{L_{BC}\beta^2}{2cos\bar{\theta}} - \frac{L_{BC}\beta}{sin\bar{\theta}} + \frac{L_{AD}\alpha}{sin\bar{\theta}}\right] \\ \Delta\theta_1 \approx \frac{1}{2L_{AB}}\left[\frac{L_{AD}\alpha^2}{2cos\bar{\theta}} - \frac{L_{BC}\beta^2}{2cos\bar{\theta}} - \frac{L_{AD}\alpha}{sin\bar{\theta}} + \frac{L_{BC}\beta}{sin\bar{\theta}}\right] \end{cases} \tag{A17}$$

Obviously, the derivative of x and y can be achieved from Equations (A16) and (A17):

$$\begin{cases} x' = \left(L_0 + \frac{cos\bar{\theta}}{2sin\bar{\theta}} L_{\mathrm{AD}}\right)\alpha' - \left(L_G + \frac{cos\bar{\theta}}{2sin\bar{\theta}} L_{\mathrm{BC}}\right)\beta' \\ x'' = \left(L_0 + \frac{cos\bar{\theta}}{2sin\bar{\theta}} L_{\mathrm{AD}}\right)\alpha'' - \left(L_G + \frac{cos\bar{\theta}}{2sin\bar{\theta}} L_{\mathrm{BC}}\right)\beta'' \end{cases} \tag{A18}$$

$$\begin{cases} y' = \left(\frac{cos\bar{\theta}}{2sin\bar{\theta}} L_{\mathrm{AD}} - L_0\right)\alpha\alpha' - \left(\frac{cos\bar{\theta}}{2sin\bar{\theta}} L_{\mathrm{BC}} - L_{\mathrm{G}}\right)\beta\beta' \\ y'' = \left(\frac{cos\bar{\theta}}{2sin\bar{\theta}} L_{\mathrm{AD}} - L_0\right)(\alpha'\alpha' + \alpha\alpha'') - \left(\frac{cos\bar{\theta}}{2sin\bar{\theta}} L_{\mathrm{BC}} - L_{\mathrm{G}}\right)(\beta'\beta' + \beta\beta'') \end{cases} \tag{A19}$$

where $\Delta\theta_1$ and $\Delta\theta_2$ are the variation quantity of θ_1 and θ_2, respectively, rad.

According to the dynamic equilibrium equation,

$$Mx'' = T_1 \sin\theta_3 - T_2 \sin\theta_4 + F_1 \sin\theta_1 + F_2 \sin\theta_2 \tag{A20}$$

$$My'' = T_1 \cos\theta_3 - T_2 \cos\theta_4 + F_1 \cos\theta_1 + F_2 \cos\theta_2 - Mg \tag{A21}$$

$$I_B\beta'' = T_1\frac{L_{MN}}{2}\cos\theta_7 + T_2\frac{L_{MN}}{2}\cos\theta_8 + F_1\phi_1 - F_2\phi_2 \tag{A22}$$

$$\phi_1 = L_G \cos\theta_5 + \frac{L_{BC}}{2}\sin\theta_5 \tag{A23}$$

$$\phi_2 = L_G \cos\theta_6 + \frac{L_{BC}}{2}\sin\theta_6 \tag{A24}$$

F_1 and F_2 are the forces of link rods AB and CD, respectively, N; T_1 and T_2 are equivalent damping forces of the left-side and right-side balance cylinders, N; M boom mass, kg; I_B Moment of inertia of boom about mass center, kg·m^2; θ_5 and θ_6—intersection angles between connecting rod AB, CD and boom, respectively, rad.

It can be found that the transfer function of the hydraulic control valve and piston can be obtained by Equation (A5)

$$\frac{x_p(s)}{x_v(s)} = \frac{k_1}{s(k_2 s^2 + k_3 s + 1)} \tag{A25}$$

where $k_1 = \frac{K_q}{A_p}$, $k_2 = \frac{1}{\omega_h^2}$, $k_3 = 2\zeta_h/\omega_h$.

Assume that the hydraulic oil is incompressible and sufficient hydraulic pressure enables the piston to move independent of its load, thus $k_2 = 0$ and $k_3 = 0$ [9].

$$\frac{x_v(s)}{\gamma - \beta} = \frac{k_4}{Ts + 1} \tag{A26}$$

where k_4 is the gain of valve and transfer function, m/rad.

The transfer function of inclination transducer mounted on the boom to the hydraulic valve displacement can be written as [7]

$$\frac{x_p}{\gamma - \beta} = \frac{K}{s(Ts+1)} \tag{A27}$$

Considering that the increment in length of ME or MF is equal to the piston displacement increment ΔL_b, and thus the overall transfer function for transducer, the valve and piston combination can be rewritten as

$$G(s) = \frac{\Delta L_b}{\gamma - \beta} = \frac{K}{s(Ts+1)} = \frac{B}{s^2 + As} \tag{A28}$$

where $K = k_1 k_4$, $B = K/T$; $A = 1/T$; γ inclination of ground to horizontal (rad); β inclination of boom to horizontal (rad); T time constant of valve and transducer transfer function; s Laplace operator; k_1, k_2, k_3 coefficients of valve and piston transfer function, Equation (A24); k_1 gain of valve and transducer transfer function.

The behavior of the active suspension can be predicted by the simultaneous solutions of Equations (A13)–(A24) and (A28).

References

1. Zhuang, T.F.; Yang, X.J.; Dong, X.; Zhang, T.; Yan, H.R.; Sun, X. Research status and development trend of large self-propelled sprayer booms. *Trans. Chin. Soc. Agric. Mach.* **2018**, *49*, 189–198. (In Chinese)
2. Visacki, V.; Sedlar, A.; Gil, E.; Bugarin, R.; Turan, J.; Janic, T.; Burg, P. Effects of sprayer boom height and operating pressure on the spray uniformity and distribution model development. *Appl. Eng. Agric.* **2016**, *32*, 341–346.
3. Balsari, P.; Gil, E.; Marucco, P.; van de Zande, J.C.; Nuyttens, D.; Herbst, A.; Gallart, M. Field-crop-sprayer potential drift measured using test bench: Effects of boom height and nozzle type. *Biosyst. Eng.* **2017**, *154*, 3–13. [CrossRef]
4. Cui, L.F.; Xue, X.Y.; Le, F.X.; Mao, H.P.; Ding, S.M. Design and experiment of electro-hydraulic active suspension for controlling the rolling motion of spray boom. *Int. J. Agric. Biol. Eng.* **2019**, *12*, 72–81. [CrossRef]
5. Yan, J.C.; Xue, X.Y.; Cui, L.F.; Ding, S.M.; Gu, W.; Le, F.X. Analysis of dynamic behavior of spray boom under step excitation. *Appl. Sci.* **2021**, *11*, 14–24. [CrossRef]
6. Sun, W.F.; He, Y.; Fu, T.P.; Wang, J.; Lu, J.Q.; Chang, J.K. Design and test of horizontal folding spray boom of sprayer. *Trans. Chin. Soc. Agric. Mach.* **2021**, *53*, 116–127.
7. Manea, D.; Gidea, M.; Marin, E.; Mateescu, M. Simulation of mechanical parameters of sprayer boom. In Proceedings of the 17th International Scientific Conference on Engineering for Rural Development, Jelgava, Latvia, 23–25 May 2018; pp. 45–51.
8. Lebeau, F.; El Bahir, L.; Destain, M.F.; Kinnaert, M.; Hanus, R. Improvement of spray deposit homogeneity using a PWM spray controller to compensate horizontal boom speed variations. *Comput. Electron. Agric.* **2004**, *43*, 149–161. [CrossRef]
9. Satow, T.; Miyamoto, K.; Matsuda, K. Control of spraying height with ultrasonic sensor for boom sprayer (part 2)-Development of automatic control device for spraying height. *J. Jpn. Soc. Agric. Mach.* **1994**, *56*, 59–67.
10. Anthonis, J.; Ramon, H.; Baerdemaeker, J.D. Implementation of an active horizontal suspension on a spray boom. *Trans. ASAE* **2000**, *43*, 213–220. [CrossRef]
11. Deprez, K.; Anthonis, J.; Roman, H.; Van Brussel, H. Development of a slow active suspension for stabilizing the roll of spray booms, part 1: Hybrid Modelling. *Biosyst. Eng.* **2002**, *81*, 185–191. [CrossRef]
12. Nuyttens, D.; Zwertvaegher, I.K.; Dekeyser, D. Spray drift assessment of different application techniques using a drift test bench an comparison with other assessment methods. *Biosyst. Eng.* **2016**, *154*, 14–24. [CrossRef]
13. Cui, L.F.; Xue, X.Y.; Ding, S.M.; Le, F.X. Development of a DSP-based electronic control system for the active spray boom suspension. *Comput. Electron. Agric.* **2019**, *166*, 105024. [CrossRef]
14. Tan, H.R.; Dou, H.J.; Zhai, C.Y.; Li, Y.K.; Yang, S.; Chen, L.P. Design and test of boom height control system for boom sprayer. *J. Agric. Mech. Res.* **2021**, *6*, 156–160.
15. Dou, H.J.; Zhai, C.Y.; Chen, L.P.; Wang, S.L.; Wang, X. Field variation characteristics of sprayer boom height using a newly designed boom height detection system. *IEEE Access* **2021**, *9*, 17148–17160. [CrossRef]
16. Tahmasebi, M.; Mailah, M.; Gohari, M.; Abd Rahman, R. Vibration suppression of sprayer boom structure using active torque control and iterative learning. Part I: Modelling and control via simulation. *J. Vib. Control* **2018**, *24*, 4689–4699. [CrossRef]
17. Andrea, W.; Miklos, K. State feedback controller design of an active suspension system for vehicles using pole placement techenique. *Acta Tech. Jaurinensis* **2019**, *12*, 178–190.
18. Xue, T.; Li, W.; Du, Y.F.; Mao, E.R.; Wen, H.J. Adaptive fuzzy sliding mode control of spray boom active suspension for large high clearance sprayer. *Trans. CSAE* **2018**, *34*, 47–56. (In Chinese)
19. Anthonis, J.; Ramon, H. Design of an active suspension to suppress the horizontal vibrations of a spray boom. *J. Sound Vib.* **2003**, *266*, 573–583. [CrossRef]

20. Available online: https://amazone.co.uk/ (accessed on 8 November 2023).
21. Available online: https://lemken.com/en-en/# (accessed on 6 November 2023).

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.

agriculture

MDPI

Article

Embedded Field Stalk Detection Algorithm for Digging–Pulling Cassava Harvester Intelligent Clamping and Pulling Device

Wang Yang *, Junhui Xi, Zhihao Wang, Zhiheng Lu, Xian Zheng, Debang Zhang and Yu Huang

College of Mechanical Engineering, Guangxi University, Nanning 530004, China; junhui_xi@163.com (J.X.); wzh0186@163.com (Z.W.); 8812316@163.com (Z.L.); xian_zheng@163.com (X.Z.); zdbang1233@163.com (D.Z.); huangyu202305@163.com (Y.H.)

* Correspondence: yanghope@163.com

Abstract: Cassava (*Manihot esculenta* Crantz) is a major tuber crop worldwide, but its mechanized harvesting is inefficient. The digging–pulling cassava harvester is the primary development direction of the cassava harvester. However, the harvester clamping–pulling mechanism cannot automatically adjust its position relative to the stalks in forward movement, which results in clamping stalks with a large off-center distance difficulty, causing large harvest losses. Thus, solving the device's clamping location problem is the key to loss reduction in the harvester. To this end, this paper proposes a real-time detection method for field stalks based on YOLOv4. First, K-means clustering is applied to improve the consistency of cassava stalk detection boxes. Next, the improved YOLOv4 network's backbone is replaced with MobileNetV2 + CA, resulting in the KMC-YOLO network. Then, the proposed model's validity is demonstrated using ablation studies and comparison tests. Finally, the improved network is embedded into the NVIDIA Jetson AGX Xavier, and the model is accelerated using TensorRT, before conducting field trials. The results indicate that the KMC-YOLO achieves average precision (AP) values of 98.2%, with detection speeds of 33.6 fps. The model size is reduced by 53.08% compared with the original YOLOv4 model. The detection speed after TensorRT acceleration is 39.3 fps, which is 83.64% faster than before acceleration. Field experiments show that the embedded model detects more than 95% of the time at all three harvest illumination levels. This research contributes significantly to the development of cassava harvesters with intelligent harvesting operations.

Keywords: digging–pulling cassava harvester; intelligent clamping; stalk section detection; YOLOv4 optimization algorithm; embedded platform

Citation: Yang, W.; Xi, J.; Wang, Z.; Lu, Z.; Zheng, X.; Zhang, D.; Huang, Y. Embedded Field Stalk Detection Algorithm for Digging–Pulling Cassava Harvester Intelligent Clamping and Pulling Device. *Agriculture* **2023**, *13*, 2144. https://doi.org/10.3390/agriculture13112144

Academic Editor: Jin He

Received: 10 October 2023
Revised: 22 October 2023
Accepted: 30 October 2023
Published: 14 November 2023



Copyright: © 2023 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (https://creativecommons.org/licenses/by/4.0/).

1. Introduction

Cassava (*Manihot esculenta* Crantz) is one of the world's three major crops, the third largest food crop in the tropical region, the world's sixth largest food crop, and 1 billion people's food rations, known as "the king of starch". It is widely planted in more than 100 countries or regions in Asia, Africa, and Latin America and also an important energy crop and industrial raw material [1–3]. However, cassava harvesting primarily relies on manual methods, where farmers either pull up the tubers by hand or use a type of manual lever harvester [4]. This heavy reliance on manual labor for harvesting severely restricts the advancement of the cassava industry. Consequently, there is a critical need for research and development in cassava harvesting machinery.

Currently, three types of cassava harvesting machines are available [5]. The first type is the digging type of cassava harvester, which includes the 4UMS390II cassava harvester designed by Xue et al. [6], the TEK cassava harvester in Ghana [7], and the cassava harvester model P900 in Brazil [8], as shown in Figure 1a. This type of harvester can only function at semi-mechanized operation and requires manually pulling out the tubers after loosening

the soil. Although they exhibit lower operational efficiency, they are well suited for various soil conditions. The second type is the digging and shaking separation type of cassava harvester, which includes the vibrating cassava root harvester developed by Gupta. et al. [9] in Thailand, the classic model API cassava harvester developed in Malaysia [10], the 4UMZ-1400 rear-collected cassava combine harvester designed by Li et al. [11], and the 4UM-160 cassava harvester developed by Mo and Huang [12], which adopts the technical route of shoveling (cutting), soil breaking, soil sieving, separating–lifting, and placing, as shown in Figure 1b. This kind of harvester can function at the fully mechanized level, with great operating efficiency, but it consumes a lot of power and is poorly adapted to diverse soil types where cassava is cultivated, with less tuber loss and damage in sandy soils and more tuber loss and damage in clayey soils. The third type is the digging–pulling cassava harvester, which includes the cassava root digging–pulling type of harvester in Cuba [5], the cassava harvester developed by the University of Leipzig in Germany [13], and a series of digging–pulling types of cassava harvesters developed by Hainan University [14] and Guangxi University [15], as shown in Figure 1c, among others. This kind of harvester can achieve fully mechanized operation, relatively high operational effectiveness, low power consumption, and excellent soil adaptability. Therefore, the primary direction of cassava harvester development has been the digging–pulling type.

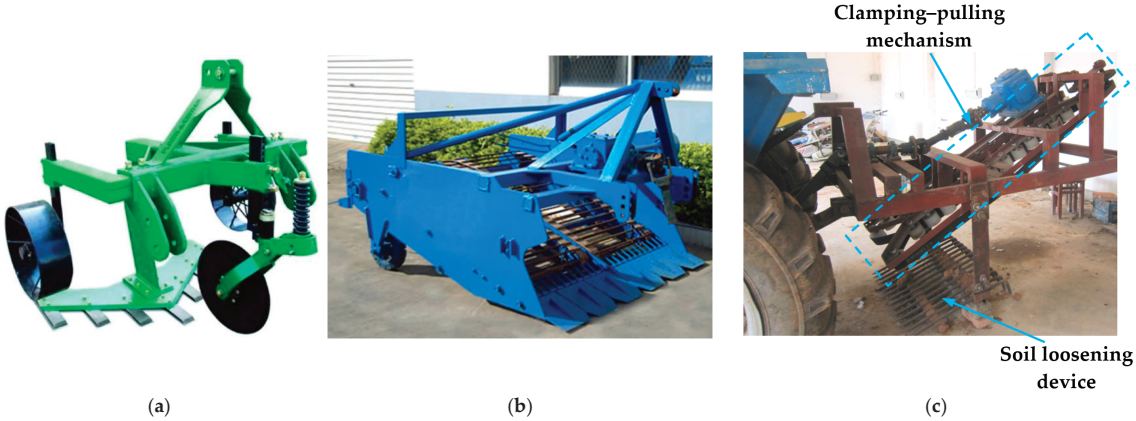


Figure 1. Typical models in three types of mainstream cassava harvesters: (**a**) digging cassava harvester; (**b**) digging and shaking separation type cassava harvester; (**c**) digging–pulling cassava harvester.

As illustrated in Figure 1c, the digging–pulling cassava harvester consists of a soil loosening device, a stalk clamping–pulling mechanism, and other components. The traditional working principle is as follows: during harvesting, the digging shovel loosens the soil first, and then, the clamping–pulling mechanism, located directly above it, clamps the stalks, pulls up the tubers, and conveys the tubers to the upper rear. Since cassava stalks present a thin and tall condition, the stalks tend to grow at an incline due to wind force, resulting in deviation from the vertical direction, as shown in Figure 2. The off-center distance is defined in this paper as the distance of the cross-section of cassava stalks 30 cm above the ground at harvest from the centerline between the two furrows, as illustrated in Figure 3. According to a significant number of field statistics, 84.5% of the stalks had an off-center distance of less than 200 mm, 9.3% had an off-center distance of 200–300 mm, and 6.2% had an off-center distance of more than 300 mm [16]. However, the existing digging–pulling harvester's clamping–pulling mechanism cannot automatically adjust its position relative to the stalks in forward movement, which results in clamping stalks with large off-center distance difficulty, causing large harvest losses. Thus, solving the device's clamping location problem is the key to loss reduction in the harvester. But the traditional

sensor fusion approach [17] is difficult to apply to stalk clamping with a considerable off-center distance. As a result, in this work, stalks are recognized during the harvesting stage by carrying an embedded detection platform to gather stalk position information relative to the harvester. These data will be utilized to direct the automatic adjustment of the movement of the clamping–pulling mechanism, which will improve stalk clamping success and reduce harvesting loss.



Figure 2. The growth of cassava stalks.

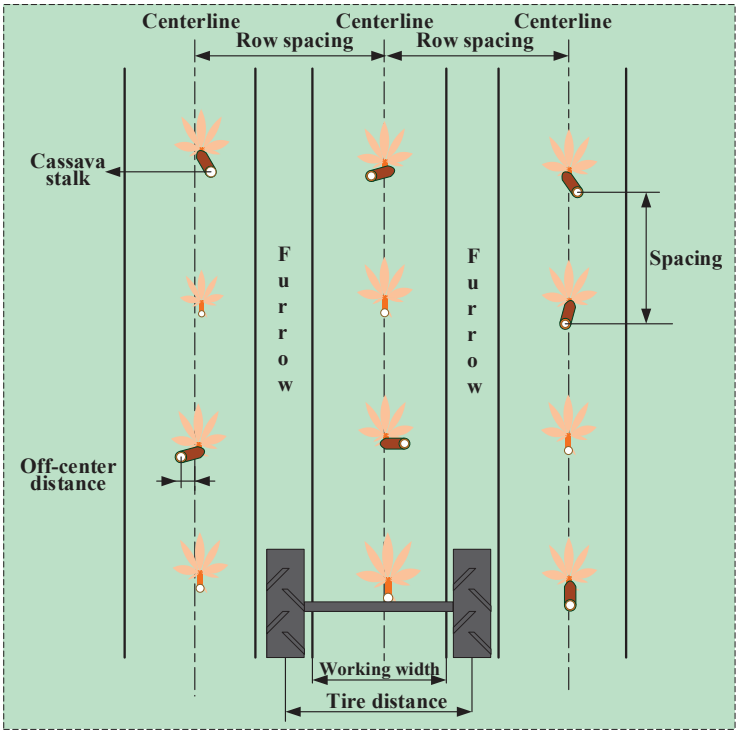


Figure 3. Cassava planting patterns and cassava stalk distribution during harvest.

Nowadays, object detection has become an effective means to assist in the positioning of mechanical structures, which can help the digging–pulling cassava harvester obtain an accurate position of the stalk for clamping. Different object detection algorithms have been successively applied to the detection and positioning of agricultural products by

researchers [18] and have achieved certain results. Among the two-stage detection algorithms, the representative Faster R-CNN algorithm has been used for fruit detection and has achieved great detection results [19]. Manual kiwifruit harvesting in orchards is labor-intensive, and Song et al. [20] have developed a VGG16-based Faster R-CNN all-weather working harvesting robotic vision system with good detection of kiwifruit images collected under different lighting conditions (morning, afternoon, and evening). A robotic vision system for multi-class fruit detection based on Faster R-CNN has been proposed by Wan and Goudos [21], who have established a library of outdoor fruit images and optimized the structure of the convolution and pooling layers in the model, and the mAP (mean Average Precision) of apples, mangoes, and oranges was above 91%. In the one-stage detection algorithm, Peng et al. [22] replaced the VGG16 network of the SSD algorithm with the ResNet-101 network for fruit detection, achieving an average accuracy of 88.4%. Gai et al. [23] improved the YOLOv4 model for cherry fruit detection by changing the prior bounding box to a circular marker box in line with the fruit shape and replacing the backbone network with the DenseNet network model, achieving a higher mAP value than YOLOv4.

Compared to the detection of fruit, the cassava stalk section is more difficult to distinguish from the soil due to the similar color, especially during field cassava tuber harvesting when illumination values vary greatly. Therefore, fully considering the influencing factors in constructing a cassava stalk dataset is crucial for successful field detection. Quan et al. [24] proposed an improved Faster R-CNN model based on a Field Robot Platform (FRP), which used five industrial USB cameras for data acquisition to capture a large number of sample images from different shooting angles. This system provides the basis for the extraction and detection of maize seedlings at different growth stages in a complex field environment. Junos et al. [25] used an improved YOLO model to detect scattered fruits in the oil palm plantation and adopted data enhancement methods such as fuzzy enhancement to simulate the actual natural environment. The results showed an outstanding average precision of 99.76% of the UAV image, with a detection time of 34.06 ms. Overall, object detection algorithms have shown great potential for improving agricultural automation, but their effectiveness depends on various factors such as lighting conditions, image quality, and dataset construction.

Although object detection algorithms are increasingly being utilized in agricultural engineering, the working environment of agricultural machinery in the field is terrible, and the expense of establishing large-scale computing platforms on it is prohibitively expensive, with no guarantee of dependability. Embedded platforms are tiny in size and can provide greater performance and reduced power consumption, as well as greater dependability and security and field edge computing at a cheaper cost [26]. On the other hand, the computational complexity of deep learning networks remains a challenge for their practical implementation in vehicular field environments with limited computing power. To address this issue, researchers have proposed various methods to simplify the current object detection models and improve real-time detection efficiency. Simplifying the network structure is an effective method to reduce the amount of calculation. For instance, Fu et al. [27] constructed the YOLO-Banana network by simplifying the network layer, which reduced the model size and shortened the detection time, providing a broad application prospect for the intelligent management and harvest in banana orchards. Another approach is to streamline the model by using a lightweight network structure to improve detection accuracy. The MobileNet lightweight network [28] with a streamlined architecture is one of the most successful application models. This network uses a depthwise separable convolution operation instead of standard convolution, which reduces the amount of calculation several times and greatly improves the network operation speed. On this basis, the MobileNetV2 lightweight network [29] uses an inverted residual, removing ReLU to avoid information loss. In addition, the MobileNetV3 lightweight network [30] uses neural architecture search (NAS) parameters to redesign the time-consuming layer structure and introduce the attention mechanism, making the network more accurate and efficient.

For existing object detection algorithms, researchers have also proposed corresponding lightweight versions, such as Light-Head R-CNN [31], Tiny-YOLO, Tiny-SSD [32], etc.

The objective of this study is to develop a real-time cassava stalk position detection method for adjusting the movement of the clamping–pulling mechanism. Firstly, the field environment and the growth state of the cassava are analyzed to determine a real-time cassava stalk coordinate acquisition scheme. Next, an image acquisition platform is established to acquire cassava stalk images and build a dataset. Then, the YOLOv4 model is adapted to improve and simplify its structure. The accuracy and real-time performance of the improved network are compared with those of different models to meet the detection requirements. Finally, the optimal detection model is deployed into the NVIDIA Jetson AGX Xavier (hereafter referred to as Xavier) embedded device for field trial validation. The results of this paper can provide important technical support for the development of digging–pulling cassava harvesters and intelligent harvesting research.

2. Materials and Methods

2.1. Dataset Production

2.1.1. Image Acquisition

The outdoor image acquisition site is located in cassava cultivation area A, at the teaching and research base of Guangxi University College of Agriculture, Nanning, Guangxi, China (22°51′ N, 108°17′ E). Cassava was cultivated in a conventional ridged pattern with row spacing of 1 m and plant spacing of 0.7–1 m. As shown in Figure 3, the stalks grew to a height of 2–2.5 m and their diameters ranged from 25–40 mm. Like with manual cassava harvesting, stalks more than 30 cm above the ground must be manually chopped off and debris swept away before the digging–pulling cassava harvester operation. As displayed in Figure 4, the stalks that remained after truncation varied in thickness and attitude, and there were various off-center positions in the cassava stalks.



Figure 4. Situation of cassava stalks cut off before harvesting.

The field movement of harvesters is complicated. In order to collect stalk images while the harvester is moving in the field and reduce the influence of the cassava stalk postures and the complexity of detection, a field acquisition platform was constructed on a digging–pulling cassava harvester developed by Guangxi University, as shown in Figure 5. The Xavier development board (NVIDIA Corporation, Santa Clara, CA, USA), camera and lens, and related supporting equipment (as shown in Table 1) were installed on the harvester. Specifically, the industrial camera was mounted on the front end of the cassava harvester frame, and the lens was arranged vertically in line with the field to detect the top section of the cassava stalk, as shown in Figure 6. The camera was arranged before the clamping mechanism to allow enough time for the movement of the clamping–pulling mechanism. The camera was arranged vertically in line with the field so that the camera field of view and the lower plane of the frame can be roughly coincident, and a machine coordinate system was constructed in which the camera detection field of view and the frame of the harvester coincide. In this way, the harvester can more easily obtain the

position information of the cassava stalks relative to the rack, facilitating the harvester clamping device to directly obtain the position coordinates of the cassava stalks for real-time adjustment and also reducing the use of arithmetic power. During the acquisition process, the cassava harvester simulates the harvesting state by driving along the ridge in a straight line. When the collector observes that the cassava stalk cross-section is fully presented on the screen, the cassava stalk cross-section photographs are collected manually.

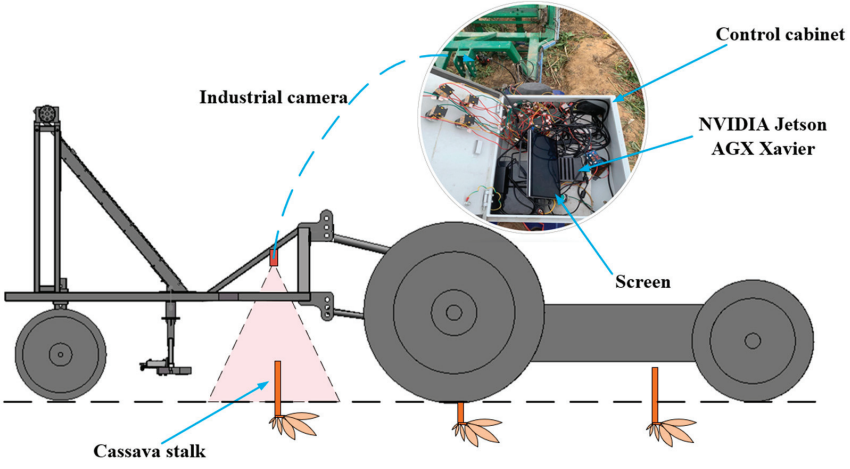


Figure 5. Field image acquisition platform.

Table 1. Image acquisition platform equipment.

Carrier	Computing Platform	Camera/Lens
Lovol M504-E tractor	NVIDIA Jetson AGX Xavier	Alvium 1800 U-508M/ KOWA LM8JC5MC



Figure 6. Cassava stalk truncated cross-section (test section).

The working principle of a digging–pulling cassava harvester equipped with an embedded inspection platform is as follows: when the cassava harvester travels along the ridge in a straight line, the camera, vertically arranged at the front end of the cassava harvester frame, acquires the positional information about the cross-section of the cassava stalks on the ridge in the field of view and transmits the information to the slave computer (Programmable Logic Controller) after processing. The slave computer controls the

movement of the clamping–pulling mechanism according to the information, and finally, the clamping–pulling mechanism accurately clamps the cassava stalks and completes the pulling and harvesting action.

In this study, images were captured using OpenCV paired with Vimba5.0, which is the official development package for Allied Vision USB3.0 interface cameras (Allied Vision Technologies Gmbh., Stuttgart, Germany). The field of view for object detection is approximately 600 mm × 500 mm, with a detection height of 600 mm. The camera has 5×10^6 pixels with a resolution of 0.25 mm per pixel. However, changes in illumination intensity during harvest can lead to changes in the color presented by the object, resulting in large differences in chromaticity information. The illumination intensity under natural conditions, such as cloudy days and evenings, ranges from 500 lux to 6000 lux, and the soil will behave in a non-granular manner. On a sunny day, when the illumination intensity is greater than or equal to 100,000 lux, the cassava stems will appear as white reflected light [33]. Therefore, to represent these variations in illumination intensity, an illumination intensity detector (Dongguan Wanchuang Electronic Products Co., Ltd., Dongguan, China, range 1–200,000 Lux, measurement accuracy (±4% rdg ± 10 dgt for ≤10,000 Lux, ±5% rdg ± 10 dgt for >20,000 Lux)) was used to detect the illumination intensity before capturing the image, as shown in Figure 7a. Then, the images can be classified accurately.



Figure 7. Detection of light intensity: (**a**) field; (**b**) indoor.

When capturing images of cassava stalk sections in the field, it is necessary not only to cut off the stalks as required but also to consider the range of illumination values for image acquisition. The crop harvest is time-sensitive and cannot be put on hold due to weather or light, and the truncated cassava stalks are very susceptible to mold after being wetted by rain. It is very difficult to acquire effective images in the field, and in order to have enough images for model training, an indoor field image dynamic simulation acquisition platform was built based on a simulated field acquisition environment, as shown in Figure 8, using a shade cloth with an adjustable LED light source to cooperate for shooting [34,35]. According to the range of illumination values, the illumination conditions of the indoor inspection platform were controlled by light intensity detection instruments as well as LED lighting devices, as shown in Figure 7b, to acquire image data under different illuminations. At the same time, the field motion condition was simulated with a panning device. A total of 1000 images were finally collected, of which 1/3 of the image data were acquired in the field and 2/3 of the cassava stalk section images were collected by the indoor acquisition platform. Finally, the illumination intensity was divided into weak light, normal light, and strong light according to the range of 500 lux to 6000 lux, 6000 lux to 100,000 lux, and greater than or equal to 100,000 lux, respectively, and the stalk side light images were additionally taken under random illumination intensity, for a total of four types of cassava stalk cross-section images, with the number of each type of image collection accounting for about one-fourth.

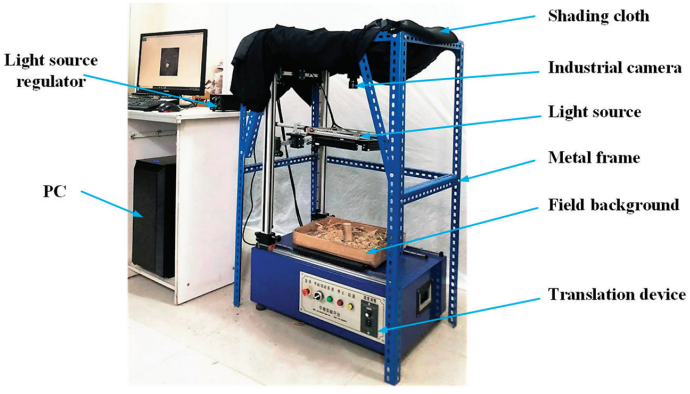


Figure 8. Indoor field image dynamic simulation acquisition platform.

2.1.2. Data Pre-Processing

This paper employed OpenCV coding to perform data augmentation on the original images so as to simulate the conditions that may be encountered in the field vehicle environment at the same time, and to improve model generalization and prevent the overfitting phenomenon. Prior to model training, the acquired raw images were randomly assigned in a 10:1 training-to-test set ratio. As illustrated in Figure 9, five data improvement methods, including motion blur, geometric changes, added noise, elastic changes, and random masking, were applied to the training and test sets, respectively. The image dataset was expanded to 11,000 pieces, comprising 10,000 images for the training set and 1000 images for the test set. As supervised learning in machine learning requires labeled datasets, GT labeling of cassava stalk cross-sections was performed using the image labeling tool LabelImg. The annotation results were stored in an XML file according to the PASCAL VOC format, which contains information such as target location, annotation box size, and category of the image, labeled as stem.

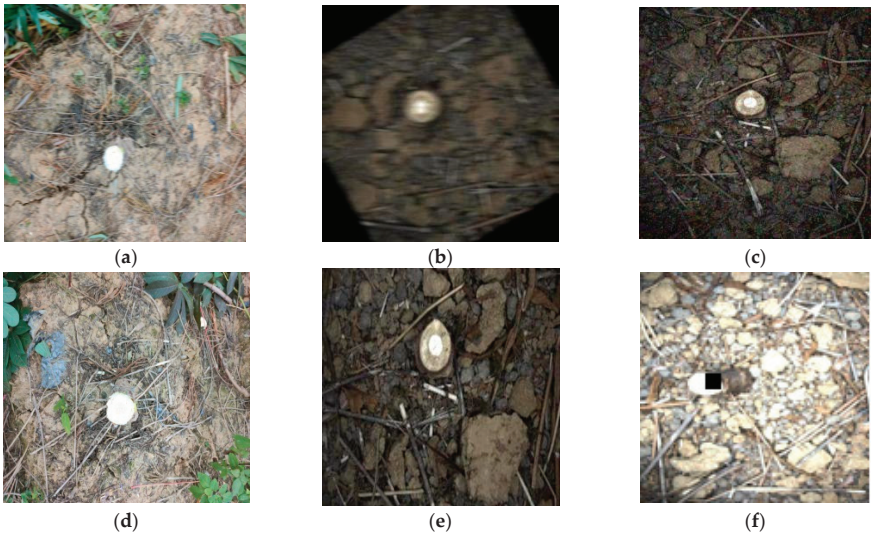


Figure 9. Data amplification method: (**a**) motion blur; (**b**) geometric rotation; (**c**) impulse noise; (**d**) Gaussian noise; (**e**) elastic transform; (**f**) random erasing (Black square is filled with zero pixel values).

2.2. Classical Model

2.2.1. YOLOv4

YOLOv4 is a typical one-stage detection algorithm that utilizes CSPDarknet-53 as the backbone network and the PANet (Path Aggregation Network) structure for multi-scale feature fusion. The network structure of YOLOv4, illustrated in Figure 6, comprises an input layer, a backbone network layer, a feature enhancement layer, and a classification regression layer. During the detection process, the image is partitioned into S×S grids, with each grid responsible for detecting the object whose center point is located within that region. Compared with the YOLOv3 network, Bochkovskiy et al. [36] incorporated the experience of CSPNet (Cross Stage Partial Network) and improved the backbone structure, which includes five CSP modules and (3 + 2 × X) convolutional layers in each CSPX, resulting in a total of 72 convolutional layers in the entire backbone. This modification addresses the issue of high forward computation in the network. Additionally, the Mish activation function [37] is used in the backbone, while the Leaky_ReLU activation function is still employed later in the network. The Mish function theoretically allows for a relatively small negative gradient inflow, thus ensuring information flow instead of an absolute zero bound as in ReLU. Moreover, the smoothed activation function enables better information penetration into the neural network, resulting in better accuracy and generalization.

2.2.2. Evaluation Metrics

For evaluating objective detection algorithms, performance evaluation metrics are necessary. The following two types of metrics are used to evaluate the models according to the evaluation metrics of neural network models.

(1) Precision, Recall, and Average Precision (AP)

The confusion matrix in Table 2 is usually used in target detection to describe the detection.

Table 2. Confusion matrix.

Confusion Matrix		**Predicted Position**	**Negative**
Actual	Positive	True Positive (TP)	True Negative (TN)
	Negative	False Positive (FP)	False Negative (FN)

Precision represents the ratio of the number of true positive samples to the total number of samples determined as positive, as shown in Equation (1). It measures the accuracy of positive predictions. Recall represents the ratio of the number of true positive samples to the number of all positive samples, as shown in Equation (2). It measures the ability of the model to retrieve all positive results. In order to balance between precision and recall, an evaluation index called average precision (AP) is commonly used to evaluate the effectiveness of object detection models. AP is calculated by computing the area under the precision–recall curve, as shown in Equation (3). In the context of cassava harvesting, the detection accuracy needs to be above 95% to ensure minimal harvesting losses and meet the requirements of the harvesting process.

$$Pre = \frac{TP}{TP + FP} \tag{1}$$

$$Recall = \frac{TP}{TP + FN} \tag{2}$$

$$AP = \int Precision \times Recall = \int_0^1 p(r)dr \tag{3}$$

(2) Detection rate FPS

FPS (Frames Per Second) indicates the number of images detected in 1 s of time and is used to represent the detection rate. The object detection network should also ensure the speed of detection with high accuracy, and this system requires the detection model to detect at 30 fps or more on the embedded development board, thus realizing real-time object detection [38].

2.3. Proposed Model

The object detection network requires high accuracy for recognizing cassava stalk sections during harvesting in a vehicle motion state. Therefore, we optimized the YOLOv4 network, comparing it to other methods and adapting it to embedded devices. First, the cassava stalk section dataset produced in this paper was clustered and analyzed using the K-means algorithm, and the obtained anchor boxes were improved through multi-scale scaling to initially improve the detection accuracy of YOLOv4. The resulting network was named K-YOLO. Second, the network model was designed to be lightweight to create an efficient network computation model. Finally, an attention mechanism was added to guarantee network performance while ensuring a lightweight network design [39].

2.3.1. Anchor Box Clustering Analysis Based on the K-Means Algorithm

The anchor box is a rectangular frame based on the common size and proportion of the object to be detected, and its similarity to the prediction box is one of the important prerequisites for accurate object prediction. The core problem of the K-means clustering algorithm is how to represent the sample-to-sample distance, and different distance calculation methods yield different clustering effects. In this paper, we used the classic IOU (Intersection over Union) calculation method of the YOLO series to calculate the distance, as shown in Equation (4):

$$d(box, centroid) = 1 - IoU(box, centroid) \tag{4}$$

where box is the bounding box and $centroid$ is the cluster center. The larger the IOU between box and the corresponding cluster center (anchor), the closer the distance.

2.3.2. Multi-Scale Scaling

The vertical distance during indoor shooting is relatively constant, and the aspect ratio and size of the stalk section are more consistent. The size of the anchor box obtained by distance-based clustering does not differ significantly across different sensing fields, which can impact the model's advantage of multi-scale output and does not align with the complex field conditions. Therefore, it is necessary to perform multi-scale random linear scaling on the obtained anchor box simultaneously in terms of aspect ratio. The scaling rules are as follows:

$$\begin{cases} x'_s = ax_s \\ x'_m = bx_m \\ x'_i = \frac{(x_i - x_s)}{(x_m - x_s)}(x'_m - x'_s) + x'_s \\ y'_i = x'_i \frac{y_i}{x_i} \end{cases} \tag{5}$$

where a and b denote the maximum and minimum scaling multipliers, x_s and x_m are the anchor frames before scaling, x'_s and x'_m denote the maximum and minimum anchor boxes, respectively.

2.3.3. Network Lightweighting

Using lightweight networks on embedded devices is a practical solution, and one such approach is to modify the K-YOLO network structure by adopting MobileNet. MobileNet is specifically designed for mobile or embedded devices. MobileNetV2 is an upgraded version of MobileNetV1, which introduces an image first before dimension raising and convolution, then the inverse residual structure of dimension reduction, and finally the

activation function of ReLU6. Moreover, it replaces the N × N matrix with the idea of 1 × N, N × 1, enabling the network to retain more low-dimensional information. MobileNetV3 further improves on MobileNetV2 by introducing the SE-Net (Squeeze and Excitation Network) attention module. However, SE-Net only considers measuring the importance of each channel by modeling channel relationships, ignoring the location information which is important for generating spatial coordinates. To solve this issue, the coordinate attention module (CA module) [40] was proposed, which considers not only the relationship between channels but also the location information in the feature space, thereby helping the model better localize and identify targets.

The K-YOLO network was improved by replacing the CSPDarknet53 backbone in the original structure with MobileNetV2 as the classification network, resulting in the KM-YOLO network. However, the streamlined model structure led to detection accuracy degradation. To solve this problem, the CA module was inserted between the global average pooling layer of the bottleneck and the 1 × 1 convolutional layer to obtain the KMC-YOLO network, as shown in Figure 10.

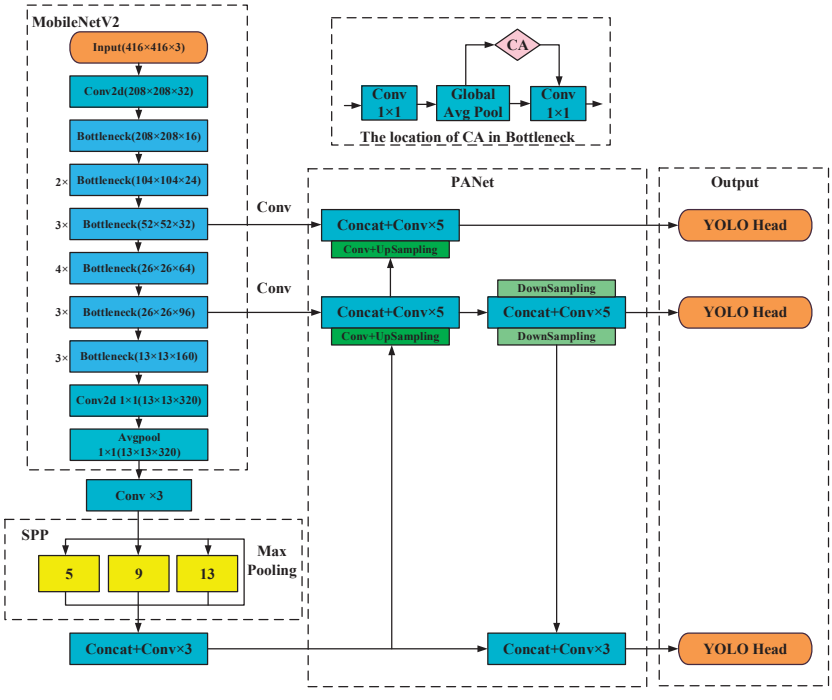


Figure 10. KMC-YOLO network structure: conv refers to convolution; SPP is spatial pyramid pooling; CA is coordinate attention; Conv is convolutional layer.

2.4. Training Platform and Environment

The object detection models were first trained and tested on a PC, using the PyTorch deep learning framework with CUDA and CUDNN installed for GPU acceleration. The training environment's composition and version are shown in Table 3.

Table 3. Algorithm operating environment.

Name	Versions
Operating System	Ubuntu18.04
CPU	Inter Core (TM) i5-7500 CPU @ 3.40 GHz, RAM:8.00 GB
GPU	NVIDIA GeForce GTX 2060-Super, RAM:8.00 GB
Compiler Environment	Pycharm
OpenCv	4.1.0.25
PyTorch	1.7.0
Python	3.8
CUDA, CUDNN	11.0, 8.05

2.5. Training Strategy

2.5.1. Training Strategy for K-Means Clustering

To accelerate the model training and enhance its generalization ability, transfer learning [41,42] was employed in this study, whereby pre-trained weights of the COCO (Common Objects in Context) dataset were used as the backbone network during YOLOv4 training. The training was conducted for a total of 200 epochs. The initial learning rate was 0.0001; the learning rate at the end of training was 0.00001. The followings values were also obtained: a weight decay value of 0.0005, a warm-up parameter value of 2 epochs, and a batch size of 4. Optimization of the parameters was achieved using the small batch gradient descent method with Momentum = 0.9.

2.5.2. Training Strategy for Lightweight Networks

The KMC-YOLO network was a redesigned network structure without a corresponding pre-training model; thus, it was trained from scratch to update the weights without using transfer learning. During training, multi-scale training was enabled, warm-up was set to 2 epochs, and the weight decay value was adjusted to 0.0003 to accommodate the simpler network structure.

2.6. Ablation Study and Comparison Test

Ablation studies were built using the YOLOv4 model with the adoption of five network models to validate the efficacy of the suggested method. Instead of K-YOLO's backbone network, CSPDarknet53, the validation network in this research was developed using MobileNetV1, MobileNetV2, MobilnetV3, and MobileNetV3-small networks with the SE-Net attention module, respectively. The performance was then compared to that of the KMC-YOLO network.

Meanwhile, in order to provide a more comprehensive comparison, we conducted detection performance tests between our improved networks and existing mainstream networks and their improved versions. The compared models include YOLOv3-SPP, YOLOv4-tiny, YOLOX-tiny, YOLOv5s, and Faster R-CNN.

2.7. Model Deployment

The optimal network model was obtained after training and comparative validation of the network model on the PC side, but it needed to be deployed on the development board for practical applications. In this study, we used the Python environment manager Archiconda3_0.2.3 on Xavier for Python package installation and virtual environment creation [43]; the Xavier environment configuration is shown in Figure 11.

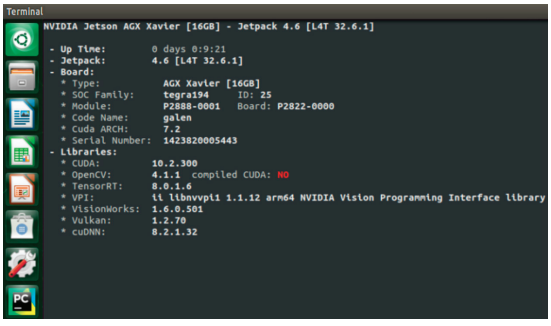


Figure 11. Development environment configuration.

The computing power of the embedded development board is limited, and running the trained deep learning model on a specific learning framework may lead to inefficient inference within the board. To overcome this issue, we used Xavier's TensorRT 8.0.1.6 to deploy the deep learning model on the development board. The implementation involved converting the PyTorch-trained model into the ONNX (Open Neural Network Exchange) intermediate format, which supports inference based on any framework, and then transforming it into the TensorRT engine based on the ONNX file format. Finally, efficient inference on the development board was achieved using the TensorRT engine file with FP16 floating-point precision.

3. Experiment Results and Discussion

3.1. Analysis of Training Results

3.1.1. Representation of Multi-Scale Anchor Boxes

The cassava dataset was clustered with the "1-IOU" distance to obtain three sets of anchor boxes (IOU anchors), and the obtained IOU anchors were randomly scaled at multiple scales to obtain three new sets of anchor boxes (CIOU anchors), as shown in Table 4. The COCO anchors are the three sets of anchor boxes obtained from the original YOLOv4 model based on the COCO dataset.

Table 4. Anchor clustering results.

Receptive Field	COCO Anchors	IOU Anchors	CIOU Anchors
Large	(116 × 90), (156 × 198), (373 × 326)	(78 × 77), (72 × 72), (75 × 61)	(140 × 114), (124 × 124), (156 × 154)
Medium	(36 × 75), (76 × 55), (72 × 146)	(65 × 67), (62 × 61), (64 × 50)	(83 × 65), (72 × 71), (88 × 91)
Small	(12 × 16), (19 × 36), (40 × 28)	(53 × 43), (48 × 50), (55 × 55)	(26 × 21), (35 × 48), (36 × 36)

The model's detection performance was tested on a test set comprising 1000 images. Table 5 shows the mAP scores for the three sets of models. The results indicate that using K-means clustering to obtain IOU anchors for YOLOv4 improves the detection accuracy compared to using COCO anchors. Moreover, using multi-scale scaling to obtain CIOU anchors further improves detection accuracy. Among the models, the K-YOLO model using CIOU anchors achieved the highest mAP score of 95.6% in the test set. With GPU acceleration, the K-YOLO model using CIOU anchors achieved a detection speed of 19.23 fps.

Table 5. Model accuracy under different anchor boxes.

Anchors	mAP (%)
COCO anchors	92%
IOU anchors	93.5%
CIOU anchors	95.6%

3.1.2. Representation of Improved Network

According to the training strategies, each of the two improved lightweight networks was trained for 200 epochs. The changes in loss values during the training of the KMC-YOLO network are shown in Figure 12.

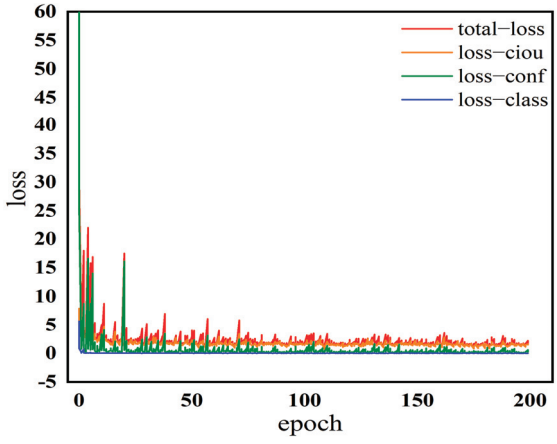


Figure 12. KMC-YOLO training process: the term total−loss refers to the overall loss, loss−ciou represents the CIoU loss, loss−conf represents the confidence loss, and loss−class represents the classification loss.

From Figure 12, it can be observed that the network loss values converge rapidly and decrease within 25 epochs. There are some small fluctuations in the confidence loss before 70 epochs, after which the fluctuations of each loss value become smaller and the convergence effect improves. The reduction in loss values for the newly designed model is quick and effective, indicating that the network structure design and hyperparameter settings are reasonable. Testing on a test set with 1000 images also shows that no overfitting occurs.

3.1.3. Validation of the Network Model

Analysis of KMC-YOLO Test Results

The test results of the KMC-YOLO network are further analyzed, and some of the results are shown in Figure 13. From the figure, it can be seen that the model demonstrates good detection performance on images captured under three different lighting conditions: bright, dark, and side light. The model also performs well on images with added processing effects such as motion blur, rotation, and distortion, including images with tilted stalks.



Figure 13. KMC-YOLO test results.

In general, the KMC-YOLO performed excellently, achieving better detection of cassava stalks under different conditions during the network test.

Results and Analysis of Ablation Studies and Comparison Tests

The upper section of Table 6 displays the results of the ablation studies, while the lower section displays the results of the comparison tests. According to the ablation studies, the size of the KM1-YOLO model is significantly reduced after replacing the backbone network with MobileNetV1, compared to the YOLOv4 model after K-means clustering (K-YOLO). This demonstrates that replacing the YOLOv4 backbone structure with the MobileNet network can achieve network lightweight, but the accuracy reduces by 4.4 percentage points. Compared to the prior model, the accuracy of the network model has been enhanced after being replaced with MobileNetV2, although the model size has more than doubled. The accuracy of the MobilnetV3 model with the self-contained SE attention module improves with small model size growth, indicating that the adoption of the attention module can improve model accuracy with little model size growth. The use of the MobilnetV3-small model with its own SE module sacrifices detection accuracy, even though the model size can be reduced relative to KM3-YOLO. Despite being able to employ a smaller model than KM3-YOLO, the MobilnetV3-small model with its own SE module reduces detection accuracy. In comparison to the other sets of models, the KMC-YOLO model suggested in this study, which combines the MobileNetV2 model and CA module, has a bigger accuracy improvement as well as improved model size and detection speed.

From the comparison test, it can be seen that the network detection performance of KMC-YOLO is better than the YOLOv3-SPP network and Faster R-CNN network in terms of detection accuracy and detection speed. Its detection accuracy is substantially higher than that of the YOLOv4-tiny, YOLOX-tiny, and YOLOv5s, although its detection speed is not as fast. The harvester clamping and pulling mechanism requires a high stalk–field detection accuracy to offer accurate stalk location and a short detection speed to allow the machine to make changes. The current model has excellent detection accuracy as well as detection speed to match the time requirements of the machinery, allowing it to be used in field detection.

Table 6. Comparison of network detection performance.

Network Model	Backbone	Precision/%	Model Size/M	Detection Speed under GPU/fps
K-YOLO	CSPDarknet53	95.6	245.3	19.23
KM1-YOLO	MobileNetV1	91.2	51.1	30.7
KM2-YOLO	MobileNetV2	93.1	113.2	33.4
KM3-YOLO	MobileNetV3	94.8	114.2	28.5
KM3-YOLO (small)	MobileNetV3-small	94.0	110.6	34.5
KMC-YOLO	**MobileNetV2 + CA**	**98.2**	**115.1**	**33.6**
YOLOv3-SPP	Darknet53	86.2	71.6	25.3
YOLOv4-tiny	CSPDarknet53-tiny	71.4	22.4	48.5
YOLOX-tiny	CSPDarknet53-tiny	86.43	19.4	48.6
YOLOv5s	CSP + Focus	89.5	14.8	52.4
Faster R-CNN	ResNet50 + FPN	84.6	137	18.5

3.2. Field Validation Trials

The detection speed of the KMC-YOLO network deployed directly into the Xavier development board is 21.4 fps with 98.1% precision. After being accelerated by TensorRT with FP16 floating-point numbers, the detection speed of KMC-YOLO is 39.3 fps with 96.8% precision, meeting the requirements for field detection. Real-time detection experiments were conducted on cassava stalk sections in the field to verify the model's landing performance after the development of the detection interface was completed. After truncation of cassava stalks in the field, the cross-section of the stalks undergo oxidization and blackening within a certain period of time, which cannot support long-term and repeated tests, so a video stream shot in cassava planting area B of Guangxi University's College of Agriculture, a teaching and research base, was used for the test. The video was shot on 11 January 2022, and the field environment information is shown in Table 7. The cassava stalks were truncated before shooting.

Table 7. Field environment information.

Weather	Environmental Temperature	Relative Humidity	Average Surface Soil Moisture Content	Cassava Variety
Sunny	8–15 °C	71%	17.22%	GR891, Bread Cassava No.1

Three different time periods were chosen for shooting: 6:30–7:00 a.m., 11:00–12:00 p.m., and 6:00–6:30 p.m. The illumination value during noon was above 10,000 lux, while the illumination value during the early morning and evening was around 1000 lux.

The shooting equipment was fixed on the cassava harvester, and the cassava harvester moved in the same way as during normal operation. The videos captured during the three time periods were saved separately. Videos of the detection process were recorded at three different times and three frames were extracted from each video. Figure 14 displays the extracted images during the test, and all three time periods yielded good recognition results. The overall detection success rate exceeded 95% in all three cases, as presented in Table 8. Upon analysis, false detections were observed in cases of strong light reflection during high noon when illumination values were high. Conversely, missed detections occurred during the morning and evening when illumination values were low. In the field test, the high illuminance value at noon was identical to the simulated environment of the indoor field image dynamic simulation acquisition platform, resulting in robust identification results and even misclassifying clods with high light backlighting as cassava stalks. The misdetection was caused by the low illuminance value in the morning and evening, as well as the gap between natural light and stabilized incandescent light in the interior platform.



Figure 14. Images of the testing process: (**a**) low light detection screen in the morning; (**b**) high light detection screen at noon; (**c**) low light detection screen in the evening.

Table 8. Detection statistics.

Detection Time	Number of Stalks Tested	Number of Correct Detections	Number of Non-Detects	Number of Error Detection	Success Rate/%
Morning	268	257	11	0	95.8
Noon	306	295	0	11	96.4
Evening	253	241	12	0	95.2

4. Conclusions

In this paper, a real-time cassava stalk localization detection method was proposed to adjust the movement of the clamping–pulling mechanism. The accuracy and real-time performance of the improved network were verified by comparing it with different models and conducting adaptive improvement. We also evaluated the effectiveness of the model after deploying it on an embedded device, NVIDIA Jetson AGX Xavier, through field trial validation. The following conclusions can be drawn from the comparison, evaluation, and validation:

(1) For the lightweight design of the YOLOv4 model, the KMC-YOLO network was constructed by incorporating the MobileNetV2 + CA module. The AP of the model was tested to be 98.2%, with detection speeds of 33.6 fps and model size reductions of 53.08%. The KMC-YOLO network is suitable for deployment on the Xavier development board.
(2) By deploying the KMC-YOLO network on the NVIDIA Jetson AGX Xavier with TensorRT acceleration, the detection speed of the network on the development board increased to 39.3 fps, which is 83.64% higher than the non-accelerated speed, satisfying the requirement of real-time detection in the field.
(3) The field test validation under different illumination conditions shows that the detection success rate of the model was above 95% under all illumination values tested, demonstrating that the algorithm met the detection requirements of the digging–pulling cassava harvester.

The detection method proposed in this paper meets the requirements for real-time and accurate detection in field harvesting and has significant advantages in terms of small memory usage and fast model operation. The proposed improved model can provide important technical support for the development of digging–pulling cassava harvesters and intelligent harvesting research. However, with the optimization of the algorithm, there is still room for further optimization of the model detection efficiency while maintaining a good detection effect. To accommodate the application of the target detection algorithm in this paper, we collected cassava stalk diameters and deflection angles via extensive statistical field trials, which guided the design of the clamping–pulling mechanism and its opening size. Now, we have manufactured a new digging–pulling cassava harvester. Finally, a cassava tuber harvesting test will be conducted in the field to completely check the operation of the cassava harvesting machine equipped with the upper vision device.

Author Contributions: Conceptualization, W.Y., J.X., Z.W. and Z.L.; methodology, J.X. and Z.W.; software, J.X. and Z.W.; validation, D.Z. and Y.H.; formal analysis, J.X. and Z.W.; investigation, D.Z. and Y.H.; resources, W.Y.; data curation, Z.W.; writing—original draft preparation, J.X.; writing—review and editing, W.Y., Z.L. and X.Z.; visualization, J.X.; supervision, X.Z.; project administration, W.Y. and Z.L.; funding acquisition, W.Y. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by the National Natural Science Foundation of China (Grant Nos. 32160422 and 51365005) and the Guangxi Natural Science Foundation (Grant No.2023GXNSFAA026376).

Institutional Review Board Statement: Not applicable.

Data Availability Statement: The data presented in this study are available in the main manuscript.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Brown, A.L.; Cavagnaro, T.R.; Gleadow, R.; Miller, R.E. Interactive effects of temperature and drought on cassava growth and toxicity: Implications for food security? *Glob. Chang. Biol.* **2016**, *22*, 3461–3473. [CrossRef] [PubMed]
2. Parmar, A.; Sturm, B.; Hensel, O. Crops that feed the world: Production and improvement of cassava for food, feed, and industrial uses. *Food Secur.* **2017**, *9*, 907–927. [CrossRef]
3. Li, C.; Dong, G.; Bian, M.; Liu, X.; Gong, J.; Hao, J.; Wang, W.; Li, K.; Ou, W.; Xia, T. Brewing rich2-phenylethanol beer from cassava and its producing metabolisms in yeast. *J. Sci. Food Agric.* **2021**, *101*, 4050–4058. [CrossRef] [PubMed]
4. Amponsah, S.K.; Sheriff, J.T.; Byju, G. Comparative evaluation of manual cassava harvesting techniques in Kerala, India. *Agric. Eng. Int. CIGR J.* **2014**, *16*, 41–52.
5. Yang, W.; Yang, J.; Zheng, X.T.; Jia, F.Y. Current Research and Development Trends of Cassava Root Harvest Machinery and Technology. *J. Agric. Mech. Res.* **2012**, *34*, 230–235.
6. Xue, Z.; Huang, H.; Li, M.; Di, Z.F.; Gao, W.W.; Cui, Z.D. Study on 4UMS-390 II cassava harvester. *J. Agric. Mech. Res.* **2010**, *32*, 79–81.
7. Bobobee, E.Y.H.; Yakanu, P.N.; Marenya, M.O.; Ochanda, J.P.O. Development in Ghana—Challenges, opportunities and prospects for cassava production in Africa. *J. Eng. Agric. Environ.* **2019**, *5*, 41–60.

8. Ospina, B.; Cadavid, L.F.; Garcia, M.; Alcalde, C. Mechanization of cassava production in Colombia. In *Cassava Research and Development in Asia*; Centro Internactional de Agricultura Tropical: Bangkok, Thailand, 2002; pp. 277–287.
9. Gupta, C.P.; Stevens, W.F.; Paul, S. Development of a vibrating cassava root harvester. *Agric. Mech. Asia Afr. Lat. Am.* **1999**, *30*, 51–55.
10. Akhir, H.M.; Sukra, A.B. Mechanization possibilities for cassava production in Malaysia. In *Cassava Research and Development in Asia: Exploring New Opportunities for an Ancient Crop, Proceedings of the 7th Asian Cassavaresearch Workshop, Bangkok, Thailand, 28 October—1 November 2002*; International Center for Tropical Agriculture (CIAT): Bangkok, Thailand, 2002; pp. 271–276.
11. Li, G.J.; Deng, G.R.; Wu, H.Z.; Zheng, S.; Cui, Z.D.; Huang, J. Design and experiment of 4UMZ—1400 rear-collected type cassava combined harvester. *J. Chin. Agric. Mech.* **2022**, *43*, 1–8.
12. Mo, Q.G.; Huang, M.A. Development and application of 4UM-160 cassava harvester. *Guang Xi Agric. Mech.* **2012**, *25*, 20–22.
13. Amponsah, S.K.; Addo, A.; Gangadharan, B. Review of various harvesting options for cassava. In *Cassava*; Waisundara, V.Y., Ed.; IntechOpen: London, UK, 2018; pp. 291–304.
14. Liao, Y.L.; Sun, Y.P.; Liu, S.H.; Cheng, D.P.; Wang, G.P. Development and prototype trial of digging-pulling style cassava harvester. *Trans. Chin. Soc. Agric. Eng.* **2012**, *28*, 29–35.
15. Yang, W.; Yang, R.; Li, J.; Wei, L.; Yang, J. Optimized tuber-lifting velocity model for cassava harvester design. *Adv. Mech. Eng.* **2018**, *10*, 2072049174. [CrossRef]
16. Zheng, X. Virtual Design and Simulation for Pulling Speed Controlled Cassava Harvest Machine. Master's Thesis, Guangxi University, Nanning, China, 2012.
17. Park, Y.; Son, H.I. A Sensor Fusion-Based Cutting Device Attitude Control to Improve the Accuracy of Korean Cabbage Harvesting. *J. ASABE* **2022**, *65*, 1387–1396. [CrossRef]
18. Montoya-Cavero, L.; Díaz De León Torres, R.; Gómez-Espinosa, A.; Escobedo Cabello, J.A. Vision systems for harvesting robots: Produce detection and localization. *Comput. Electron. Agric.* **2022**, *192*, 106562. [CrossRef]
19. Sa, I.; Ge, Z.; Dayoub, F.; Upcroft, B.; Perez, T.; McCool, C. DeepFruits: A Fruit Detection System Using Deep Neural Networks. *Sensors* **2016**, *16*, 1222. [CrossRef]
20. Song, Z.; Fu, L.; Wu, J.; Liu, Z.; Li, R.; Cui, Y. Kiwifruit detection in field images using Faster R-CNN with VGG16. *IFAC-PapersOnLine* **2019**, *52*, 76–81. [CrossRef]
21. Wan, S.; Goudos, S. Faster R-CNN for multi-class fruit detection using a robotic vision system. *Comput. Netw.* **2020**, *168*, 107036. [CrossRef]
22. Peng, H.X.; Huang, B.; Shao, Y.Y.; Li, Z.S.; Zhang, C.W.; Chen, Y.; Xiong, J.T. General improved SSD model for picking object recognition of multiple fruits in natural environment. *Trans. Chin. Soc. Agric. Eng. (Trans. CSAE)* **2018**, *34*, 155–162.
23. Gai, R.; Chen, N.; Yuan, H. A detection algorithm for cherry fruits based on the improved YOLO-v4 model. *Neural Comput. Appl.* **2023**, *35*, 13895–13906. [CrossRef]
24. Quan, L.; Feng, H.; Lv, Y.; Wang, Q.; Zhang, C.; Liu, J.; Yuan, Z. Maize seedling detection under different growth stages and complex field environments based on an improved Faster R–CNN. *Biosyst. Eng.* **2019**, *184*, 1–23. [CrossRef]
25. Junos, M.H.; Mohd Khairuddin, A.S.; Thannirmalai, S.; Dahari, M. Automatic detection of oil palm fruits from UAV images using an improved YOLO model. *Vis. Comput.* **2022**, *38*, 2341–2355. [CrossRef]
26. Zhang, Y.; Yu, J.; Chen, Y.; Yang, W.; Zhang, W.; He, Y. Real-time strawberry detection using deep neural networks on embedded system (rtsd-net): An edge AI application. *Comput. Electron. Agric.* **2022**, *192*, 106586. [CrossRef]
27. Fu, L.; Yang, Z.; Wu, F.; Zou, X.; Lin, J.; Cao, Y.; Duan, J. Yolo-Banana: A lightweight neural network for rapid detection of banana bunches and stalks in the natural environment. *Agronomy* **2022**, *12*, 391. [CrossRef]
28. Howard, A.G.; Zhu, M.; Chen, B.; Kalenichenko, D.; Wang, W.; Weyand, T.; Andreetto, M.; Adam, H. MobileNets: Efficient convolutional neural networks for mobile vision applications. *arXiv* **2017**, arXiv:1704.04861.
29. Sandler, M.; Howard, A.; Zhu, M.L.; Zhmoginov, A.; Chen, L.C. MobileNetV2: Inverted residuals and linear bottlenecks. In Proceedings of the 2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition, Salt Lake City, UT, USA, 18 June 2018; pp. 4510–4520.
30. Howard, A.; Sandler, M.; Chu, G.; Chen, L.; Chen, B.; Tan, M.; Wang, W.; Zhu, Y.; Pang, R.; Vasudevan, V.; et al. Searching for MobileNetV3. In Proceedings of the 2019 IEEE/CVF International Conference on Computer Vision (ICCV), Seoul, Republic of Korea, 27 October–2 November 2019; pp. 1314–1324.
31. Li, Z.; Peng, C.; Yu, G.; Zhang, X.; Deng, Y.; Sun, J. Light-Head R-CNN: In defense of two-stage object detector. *arXiv* **2017**, arXiv:1711.07264.
32. Wong, A.; Shafiee, M.J.; Li, F.; Chwyl, B. Tiny SSD: A tiny single-shot detection deep convolutional neural network for real-time embedded object detection. In Proceedings of the 2018 15th Conference on Computer and Robot Vision (CRV), Toronto, ON, Canada, 8 May 2018; pp. 95–101.
33. García-Santillán, I.D.; Montalvo, M.; Guerrero, J.M.; Pajares, G. Automatic detection of curved and straight crop rows from images in maize fields. *Biosyst. Eng.* **2017**, *156*, 61–79. [CrossRef]
34. Marx, C.; Pastrana, J.C.; Hustedt, M.; Kaierle, S.; Walter, J.; Haferkamp, H.; Rath, T. Entwicklung und Aufbau eines Lasersystem-prototyps auf Bildanalysebasis zur Unkrautbekämpfung. In Proceedings of the 17 und 18 Workshop Computer-Bildanalyse in der Landwirtschaft Computerised Image Analysis in Agriculture, Osnabrück, Germany, 9 May 2012; p. 102.

35. Zhang, B.H.; Huang, W.Q.; Li, J.B.; Zhao, C.J.; Liu, C.L.; Huang, D.F. On-line Identification of Defect on Apples Using Lightness Correction and AdaBoost Methods. *Trans. Chin. Soc. Agric. Mach.* **2014**, *45*, 221–226.
36. Bochkovskiy, A.; Wang, C.; Liao, H.M. Yolov4: Optimal speed and accuracy of object detection. *arXiv* **2020**, arXiv:2004.10934.
37. Misra, D. Mish: A self regularized non-monotonic activation function. *arXiv* **2019**, arXiv:1908.08681.
38. Choi, J.; Chun, D.; Kim, H.; Lee, H. Gaussian YOLOv3: An accurate and fast object detector using localization uncertainty for autonomous driving. In Proceedings of the 2019 IEEE/CVF International Conference on Computer Vision (ICCV), Seoul, Republic of Korea, 27 October 2019; pp. 502–511.
39. Liang, R. Research and Application of Driverless Oriented Traffic Signal Detection and Recognition Method. Master's Thesis, University of Electronic Science and Technology of China, Chengdu, China, 2021.
40. Hou, Q.; Zhou, D.; Feng, J. Coordinate attention for efficient mobile network design. In Proceedings of the 2021 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), Nashville, TN, USA, 20 January 2021; pp. 13708–13717.
41. Donahue, J.; Jia, Y.Q.; Vinyals, O.; Hoffman, J.; Zhang, N.; Tzeng, E.; Darrell, T. DeCAF: A deep convolutional activation feature for generic visual recognition. *Int. Conf. Mach. Learn.* **2014**, *32*, 647–655.
42. Song, G.H. Image Annotation Method Based on Transfer Learning and Deep Convolutional Feature. Ph.D. Thesis, Zhejiang University, Hangzhou, China, 2017.
43. Burford, A.; Calder, A.C.; Carlson, D.; Chapman, B.; Coskun, F.; Curtis, T.; Feldman, C.; Harrison, R.J.; Kang, Y.; Michalow-Icz, B.; et al. Ookami: Deployment and initial experiences. In *Practice and Experience in Advanced Research Computing*; Cornell University Library: Ithaca, NY, USA, 2021.

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.

agriculture

MDPI

Article

Motion-Control Strategy for a Heavy-Duty Transport Hexapod Robot on Rugged Agricultural Terrains

Kuo Yang [1,2], Xinhui Liu [1,2], Changyi Liu [2,3,*] and Ziwei Wang [1,2]

[1] School of Mechanical and Aerospace Engineering, Jilin University, Changchun 130025, China; yagnkuo20@mails.jlu.edu.cn (K.Y.); liuxh@jlu.edu.cn (X.L.); ziweiw20@mails.jlu.edu.cn (Z.W.)
[2] Weihai Institute for Bionics, Jilin University, Weihai 264207, China
[3] Key Laboratory of Bionic Engineering, Ministry of Education, Jilin University, Changchun 130025, China
* Correspondence: liuchangyi@jlu.edu.cn

Abstract: Legged agricultural transportation robots are efficient tools that can autonomously transport goods over agricultural terrain, and their introduction helps to improve the efficiency and quality of agricultural production. Their effectiveness depends on their adaptability to different environmental conditions, which is especially true for heavy-duty robots that exert ground forces. Therefore, this study proposes a motion-control strategy for a heavy-duty transport hexapod robot. Two critical tasks were accomplished in this paper: (1) estimating the support surface angle based on the robot's foot position and body posture, and accordingly determining the motion constraint conditions on this support surface and the body posture based on energy optimization; (2) proposing an adaptive fuzzy impedance algorithm for real-time force–position composite control for adjusting foot position, in order to reduce the steady-state force tracking error caused by terrain stiffness, thus ensuring body stability through tracking of variable foot-end forces. An element of hardware in the loop control platform for a 3.55-ton device was designed and compared with the current popular force-control methods under different external contact terrains. The results show that the proposed control method can effectively reduce force errors, establish support forces faster on less-stiff environments, and reduce the torso tilt during phase switching.

Keywords: agricultural robot; environmental adaptation; adaptive impedance control; unstructured terrain

Citation: Yang, K.; Liu, X.; Liu, C.; Wang, Z. Motion-Control Strategy for a Heavy-Duty Transport Hexapod Robot on Rugged Agricultural Terrains. *Agriculture* **2023**, *13*, 2131. https://doi.org/10.3390/agriculture13112131

Academic Editors: Jin Yuan and Zichen Huang

Received: 13 October 2023
Revised: 6 November 2023
Accepted: 9 November 2023
Published: 11 November 2023

Copyright: © 2023 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (https://creativecommons.org/licenses/by/4.0/).

1. Introduction

As robots continue to expand their role within agricultural settings, there has been extensive research on agricultural robots [1]. The primary research focus has largely been on wheeled and tracked mobile robots [2]. These robots accomplish various tasks through the integration of multiple sensor technologies such as computer vision, radar, GPS, and audio signals [3,4]. In agricultural activities, most of which are on unstructured terrain often accompanied by rugged and steep road surfaces, compared with other types of robots, the main advantage of legged robots given the unstructured terrain is their excellent terrain adaptability, together with the corresponding top-loading platforms, which can play a good role in the development of agricultural production [5]. On an unstructured terrain, studying the posture of legged robots is crucial for their adaptation to the environment. Adjusting their posture to suit the terrain can improve their adaptation ability. To achieve this, Song et al. [6] proposed a quasi-static climbing gait that enhances the movement of a robot on complex terrain. This gait enables the robot to climb steps that are 3.9 times the length of its legs. Lee [7] conducted a study on a hexapod adaptive suspension vehicle (ASV) by analyzing the robot's body posture on various inclines. A hexapod robot can traverse surfaces at different inclination levels by adjusting its torso height and posture. However, owing to the inherent variability in the end position of the foot and the susceptibility of legged robots to external environmental factors, controlling their motion on unstructured

terrain has become a subject of extensive research [8,9]. A hexapod robot can traverse surfaces at different inclination levels by adjusting its torso height and posture. Boston Dynamics [10] developed the Little Dog, a four-legged robot that optimizes its stance to maximize the movement space while ensuring that its torso does not collide with the ground. However, achieving energy efficiency is challenging for legged robots with high loads [11]. A poor distribution of foot forces can lead to energy wastage, out-of-range joint moments, and even mechanical structural damage. Therefore, rational movement is necessary to reduce energy consumption.

Impedance control [12] is a widely used robotic control technique that adjusts the force-feedback error based on the relationship between the contact force and the robot's foot-end position. Zhu [13,14] proposed a hybrid sliding film control algorithm for a hydraulic quadrupedal robot based on impedance control which resulted in improved control accuracy. Subsequently, Wang [15] developed a whole-body control architecture that included impedance control to reduce the impact of contact between the legs of a six-legged robot and the environment, thus enabling smoother contact. However, conventional impedance control models are prone to steady-state errors during force tracking owing to uncertainties associated with the environmental parameters of the terrain [16]. Several approaches have been explored to mitigate such errors, with some focusing on controlling the model reference [17–21]. For instance, Roveda [17] proposed a method to estimate the environmental stiffness and adjust the controller gain to compensate for unknown environmental stiffness, whereas Deng [18] employed an extended Kalman filter to estimate force-tracking errors. Wang [20] introduced an impedance-estimation method for estimating contact in an uncalibrated environment. However, none of these methods demonstrated satisfactory force-tracking performance under time-varying forces. Furthermore, because the aforementioned methods rely on the estimation of environmental impedance parameters, they inevitably incur errors when addressing complex and variable environments.

Researchers have explored the use of variable impedance control methods based on reinforcement learning and optimization-based approaches. Chao and Yang estimated human stiffness dynamics using a dynamic motion primitive (DMP) [22] and a Markov model with Gaussian mixed regression [23], respectively. Zhang, et al. [24] proposed a convolutional neural network (CNN) supervised-learning method to calibrate the force estimation algorithm. Modares [25] transformed the optimal parameter search problem for impedance models into a linear quadratic regulator (LQR) problem. However, these methods require extensive data collection or calculations that negatively affect the real-time tracking of the time-varying forces of a robot.

In conclusion, the current impedance-control techniques for legged robots face difficulties in achieving the desired tracking of time-varying forces in complex environments, particularly when considering real-time performance. Previous studies have demonstrated that adaptive architectures provide high robustness in uncertain dynamic systems [26–28]. Therefore, an adaptive motion strategy was proposed for complex terrains by adapting a robot torso pose to the instant terrain considering motion constraints and energy-optimal conditions under the support surface of the terrain. Subsequently, positional maintenance was achieved through adaptive fuzzy impedance control as follows:

(1) A terrain-adaptive motion strategy was proposed, one which improves the stability of the robot's motion process by estimating the angle of the support surface on the current terrain, setting up corresponding reaction mechanisms, and analyzing the energy consumption of different positions on the same support surface to select the most energy-efficient motion posture.
(2) The contact force dynamics were used to describe the influence of parameters such as environmental stiffness on the impedance model, and an adaptive fuzzy impedance control strategy based on force–position control was proposed to mitigate the impact of the external environment on impedance control.

(3) The efficacy of the proposed control strategy in complex environments was verified through comparative experiments.

This paper is organized as follows: In Section 2, we present the motion model and terrain estimation algorithm for an agricultural transportation robot with terrain adaptation on the rugged terrains of agriculture. Section 3 presents an adaptive impedance system that includes terrain-adaptive control and fuzzy parameter-adjustment methods. In Section 4, we present an experimental validation of the proposed method. Finally, Section 5 concludes the study.

2. Terrain-Adaptive Methods

The hexapod robot structure depicted in Figure 1 consists of two main components: the torso and the legs. The legs and torso's mounting positions are elliptically distributed, and each leg comprises a hip joint, thigh, and shin connected by a rotating pair between two adjacent segments, and a calf and foot-end connected by a ball hinge. An electro-hydraulic control system was selected as the transmission system, and integrated with low-impact motion trajectory planning to complete the trajectory planning. The robot can carry four people and can traverse over 1 m obstacles and across 1.2 m trenches, as well as achieve 40° longitudinal climbing and 30° lateral climbing. The robot movement includes two phases: swing and support. During the swing phase, the legs of the robot can be treated as single-leg tandem mechanisms, whereas during the support phase, the legs and the torso form parallel mechanisms. The position of the robot changes when it switches between the two phases of motion. Defined based on the established coordinate system shown in Figure 1, the coordinate transformation matrix formula for joint i can be obtained using the D-H parameter method, and the D-H parameters of the robot's single leg are shown in Table 1.

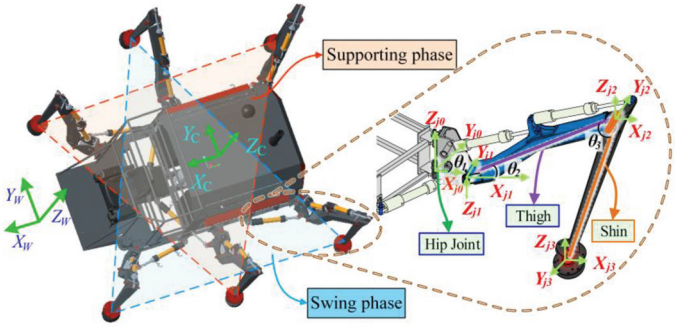


Figure 1. Structure of hexapod robot.

Table 1. D-H parameter of the leg.

	L_i	d_i	α_i	θ_i
Hip joint	l_1	0	$\pi/2$	θ_1
Thigh	l_2	0	0	θ_2
Shin	l_3	0	0	θ_3

The transformation matrix of the foot-end of leg j of the hexapod robot relative to the torso coordinate system is expressed as:

$$\begin{bmatrix} {}^{C}x_{j3} \\ {}^{C}y_{j3} \\ {}^{C}z_{j3} \end{bmatrix} = \begin{bmatrix} l_{j0} + l_1c\theta_1 + l_2c\theta_1c\theta_2 + l_3c\theta_1c(\theta_2+\theta_3) \\ b_{j0} + l_1s\theta_1 + l_2s\theta_1c\theta_2 + l_3s\theta_1c(\theta_2+\theta_3) \\ z_{j0} - l_2s\theta_2 - l_3s(\theta_2+\theta_3) \end{bmatrix} \quad (1)$$

2.1. Terrain Estimation

The overall terrain-perception process of the robot is as depicted in Figure 2. Initially, the controller detects the force signal from the foot-end force sensor, determining whether the foot-end has reached the support state based on whether the force exceeds the stable force threshold for foot grounding. Subsequently, the position of the foot-end in the support state relative to the body coordinate system is calculated through the displacement sensor of the hydraulic cylinder. The position of the supporting leg end relative to the body coordinate system is then recorded and, based on the body posture angle gathered by the IMU, the position information of the foot-end under the body coordinate system is transformed into the world coordinate system. Lastly, using the foot-end information under the world coordinate system, a plane fitting method is applied to fit the terrain information, thereby estimating the local terrain information where the hexapod robot is currently situated.

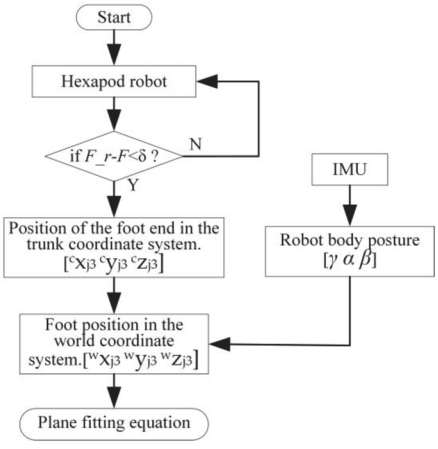


Figure 2. Flowchart of terrain awareness strategy.

To simplify the calculation of terrain estimates, a reference coordinate system, denoted by O_B-$X_BY_BZ_B$, was established to quantitatively convert the foot-end position information when characterizing the local terrain. As illustrated in Figure 3, Origin B of the reference coordinate system c coincides with Origin C of the torso coordinate system O_C-$X_CY_CZ_C$, whereas its x-, y-, and z-axis directions align with those of the reference coordinate system O_W-$X_WY_WZ_W$. The positional transformation of the torso is referenced to the world coordinate system O_B-$X_BY_BZ_B$, and the transformation matrix of the torso coordinate system with respect to the world coordinate system is given by Equation (2).

$$ {}^{B}A_{C} = \begin{bmatrix} c\gamma c\alpha & c\gamma s\alpha s\beta - s\gamma c\beta & c\gamma s\alpha c\beta + s\gamma s\beta & 0 \\ s\gamma c\alpha & s\gamma s\alpha s\beta + c\gamma c\beta & s\gamma s\alpha c\beta - c\gamma s\beta & 0 \\ -s\alpha & c\alpha s\beta & c\alpha c\beta & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \quad (2) $$

where γ, α, and β denote the yaw, pitch, and roll angles, respectively, of the torso coordinate system in the world coordinate system, as measured by the inertial measurement unit (IMU).

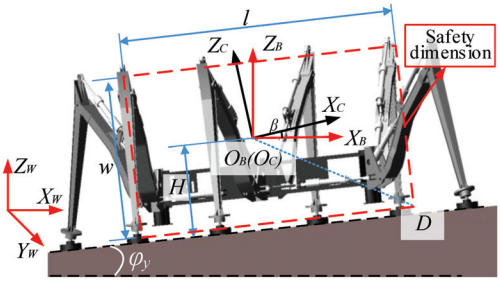


Figure 3. Diagram of reference coordinate system.

For any leg of the six-legged robot, the coordinates of the foot-end position relative to the world coordinate system can be derived using Equation (3).

$$\begin{bmatrix} {}^{B}x_{j3} \\ {}^{B}y_{j3} \\ {}^{B}z_{j3} \\ 1 \end{bmatrix} = {}^{B}A_{C} \begin{bmatrix} {}^{C}x_{j3} \\ {}^{C}y_{j3} \\ {}^{C}z_{j3} \\ 1 \end{bmatrix} \tag{3}$$

The location information of the support foot was regarded as the terrain location feature point in the terrain information, and a single-plane fitting method was employed to fit these terrain location feature points. The primary plane-fitting equation was as follows:

$$z = f(x, y) = a_0 + a_1 x + a_2 y \tag{4}$$

where a_0, a_1, and a_2 represent the coefficients of the plane equation.

The coefficients of the primary plane expression were determined using the least-squares method based on the coordinates of the foot-end position of the hexapod robot in the support phase. The solution equation is as follows:

$$\begin{bmatrix} a_0 \\ a_1 \\ a_2 \end{bmatrix} = (A^T A) A^T Z \tag{5}$$

where $A = [1\ {}^{B}x_{sj3}\ {}^{B}y_{sj3}]$, $Z = [{}^{B}z_{sj3}]$, s_{j3} is the number of legs in a supported state, and the number of legs in different gaits varies, and is obtained based on trajectory planning.

By segmenting the terrain slope based on the ground-pitch and side-inclination angles, an equation for estimating the slope of an unexplored terrain was obtained:

$$\begin{cases} \varphi_x = \arctan(a_2) \\ \varphi_y = \arctan(a_1) \end{cases} \tag{6}$$

2.2. Attitude Adaptive Adjustment

2.2.1. Torso Safety Dimension Constraints

The robot must avoid collisions between its torso and the environment when traversing complex terrain. Thus, when the terrain slope varies significantly, it is necessary to establish a geometric constraint between the robot torso safety zone dimensions and the terrain slope. The solid box in Figure 3 represents the torso profile of the robot, while the dashed line represents the minimum allowable torso safety zone dimensions for a six-legged robot with torso length l height w, and width h. The ground-pitch and ground-side-tilt angles, φ_x and φ_y, were estimated.

To ensure that the torso is not interfered with by the terrain, point D must remain above the slope throughout the motion. The torso attitude angle and geometric constraints facili-

tated the determination of the coordinates of point D relative to the reference coordinate system O_W-$X_W Y_W Z_W$.

$$\begin{cases} {}^{B}x_D = \frac{l}{2}\cos\alpha - \frac{w}{2}\sin\alpha \\ {}^{B}y_D = \frac{h}{2} \\ {}^{B}z_D = \frac{l}{2}\sin\alpha - \frac{w}{2}\cos\alpha \end{cases} \tag{7}$$

Substituting the coordinates from Equation (7) into Equation (4) yields the constraint equation for a safe torso zone size.

$$a_0 + (\frac{l}{2}c\alpha - \frac{w}{2}s\alpha)\tan\varphi_y + \frac{a_2 h}{2} - (\frac{l}{2}s\alpha - \frac{w}{2}c\alpha)\tan\varphi_x < 0 \tag{8}$$

$$a_0 = \frac{H}{\sqrt{{a_1}^2 + {a_2}^2 + 1}} \tag{9}$$

Thus, when the torso-pitch, cross-roll, ground-pitch, and ground-side-tilt angles are β, α, φ_x, and φ_y, respectively, the constraint relationship between the torso position and the terrain environment can be obtained by combining Equations (6) and (9) as follows:

$$(\frac{l}{2}c\alpha + \frac{w}{2}s\alpha s\beta + \frac{h}{2}s\alpha c\beta)\tan\varphi_y - \frac{l}{2}s\alpha - \frac{h}{2}c\alpha s\beta + \frac{H}{\sqrt{(\tan\varphi_y)^2 + (\tan\varphi_x)^2 + 1}} + (\frac{w}{2}c\beta - \frac{h}{2}c\beta)\tan\varphi_x < 0 \tag{10}$$

The robot needs to adjust its torso posture to change its direction of motion when there are obstacles in the motion's path. An RGB-D camera is installed directly in front of the robot's centroid in the x direction. The distance between the obstacle and the robot is calculated based on depth information when the target obstacle is detected. Then, by adjusting the posture of the torso, obstacle avoidance actions are performed in advance, as shown in Figure 4a. R is the safe distance to adjust the yaw angle of the torso. The calculation formula for torso deflection angle δ is as follows:

$$\delta = \arccos\frac{b_r/2 + w_2/2}{\sqrt{{x_{dr}}^2 + {y_{dr}}^2}} - \arctan\frac{y_{dr}}{x_{dr} + l} \tag{11}$$

where x_{dr} and y_{dr} are the positions of obstacles relative to camera coordinates, and their border dimensions are b_r and h_r; w_2 is the safe size of the robot's zone, and l_2 is the distance from the camera to the centroid.

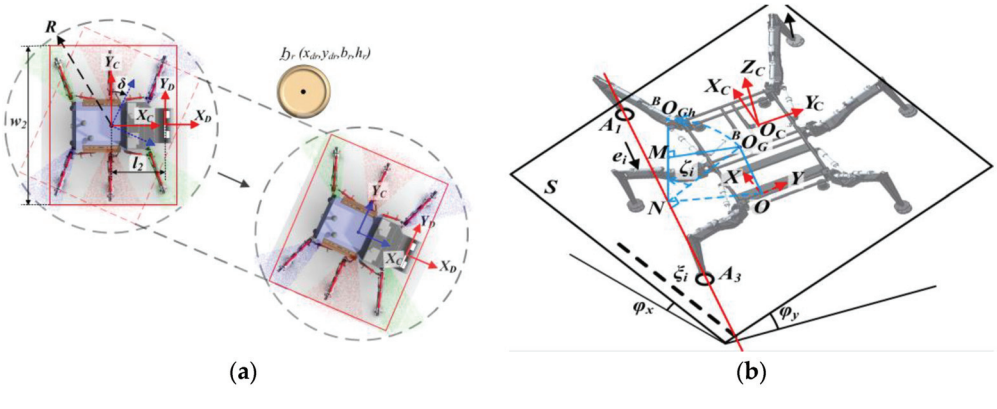


Figure 4. (**a**) Yaw angle adjustment of the torso; (**b**) sketch of the NESM solution.

2.2.2. Stability Constraints on Inclined Terrain

This study proposes a normalized energy stability margin (NESM) as a criterion for assessing the stability of a robot on an inclined terrain. The NESM method considers

the center of mass of the robot, the locations of the support feet's ends, and the slope of the terrain.

As shown in Figure 4b, the center-of-mass location in the reference coordinate system is denoted by ${}^{\mathrm{B}}O_G$ (${}^{\mathrm{B}}x_G$, ${}^{\mathrm{B}}y_G$, ${}^{\mathrm{B}}z_G$). The adjacent foot-end points of the robot for leg-end positions leg1 and leg3 are denoted as A_1 (${}^{\mathrm{B}}x_{13}$, ${}^{\mathrm{B}}y_{13}$, ${}^{\mathrm{B}}z_{13}$) and A_3 (${}^{\mathrm{B}}x_{33}$, ${}^{\mathrm{B}}y_{33}$, ${}^{\mathrm{B}}z_{33}$), respectively. The point ${}^{S}O_G$ is the projection of the center-of-mass on the slope surface S, and ${}^{\mathrm{B}}O_G$ forms a vertical line with Line $\mathrm{A}_1\mathrm{A}_3$. The intersection point of ${}^{\mathrm{B}}O_G$ and the vertical line is N (${}^{\mathrm{B}}x_N$, ${}^{\mathrm{B}}y_N$, ${}^{\mathrm{B}}z_N$). Point ${}^{\mathrm{B}}O_{\mathrm{Gh}}$ is the highest point that ${}^{\mathrm{B}}O_G$ can reach by rotating around the support edge $\mathrm{A}_1\mathrm{A}_3$. The angle between ${}^{\mathrm{B}}O_{\mathrm{Gh}}$ and the horizontal plane is denoted as θ. The angle between the support boundary c and the horizontal plane is denoted as α, and M is the intersection point of the perpendicular line from the center-of-mass ${}^{\mathrm{B}}O_G$ to the straight line ${}^{\mathrm{B}}O_{\mathrm{Gh}}N$.

The expression for the line $\mathrm{A}_1\mathrm{A}_3$ can be found from the following linear formula:

$$\frac{x - {}^{\mathrm{B}}x_{13}}{{}^{\mathrm{B}}x_{33} - {}^{\mathrm{B}}x_{13}} = \frac{y - {}^{\mathrm{B}}y_{13}}{{}^{\mathrm{B}}y_{33} - {}^{\mathrm{B}}y_{13}} = \frac{z - {}^{\mathrm{B}}z_{13}}{{}^{\mathrm{B}}z_{33} - {}^{\mathrm{B}}z_{13}} \tag{12}$$

The distance from point ${}^{\mathrm{B}}O_G$ to line $\mathrm{A}_1\mathrm{A}_3$ was calculated using the following formula:

$$D_i = \frac{\sqrt{P + K + W}}{\sqrt{Q}} \tag{13}$$

where $W = \begin{vmatrix} {}^{\mathrm{B}}z_G - {}^{\mathrm{B}}z_{13} & {}^{\mathrm{B}}x_G - {}^{\mathrm{B}}x_{13} \\ {}^{\mathrm{B}}z_{33} - {}^{\mathrm{B}}z_{13} & {}^{\mathrm{B}}x_{33} - {}^{\mathrm{B}}z_{13} \end{vmatrix}$; $K = \begin{vmatrix} {}^{\mathrm{B}}y_G - {}^{\mathrm{B}}y_{13} & {}^{\mathrm{B}}z_G - {}^{\mathrm{B}}z_{13} \\ {}^{\mathrm{B}}y_{33} - {}^{\mathrm{B}}y_{13} & {}^{\mathrm{B}}z_{33} - {}^{\mathrm{B}}z_{13} \end{vmatrix}$; and $P = \begin{vmatrix} {}^{\mathrm{B}}x_G - {}^{\mathrm{B}}x_{13} & {}^{\mathrm{B}}y_G - {}^{\mathrm{B}}y_{13} \\ {}^{\mathrm{B}}x_{33} - {}^{\mathrm{B}}x_{13} & {}^{\mathrm{B}}y_{33} - {}^{\mathrm{B}}y_{13} \end{vmatrix}$.

The normal vector of the vertical plane $\mathrm{A}_1\mathrm{A}_3\,{}^{\mathrm{B}}O_{\mathrm{Gh}}$ passing through the straight-line c is as follows:

$$n_{\mathrm{A}_1\mathrm{A}_3{}^{\mathrm{B}}O_{\mathrm{Gh}}} = \begin{bmatrix} i & j & k \\ {}^{\mathrm{B}}x_{\mathrm{G}} - {}^{\mathrm{B}}x_{13} & {}^{\mathrm{B}}y_{\mathrm{G}} - {}^{\mathrm{B}}y_{13} & {}^{\mathrm{B}}z_G - {}^{\mathrm{B}}z_{13} \\ 0 & 0 & 1 \end{bmatrix} \tag{14}$$

The vector of line ${}^{\mathrm{B}}O_{\mathrm{G}}N$ can be found when the coordinates of N and ${}^{\mathrm{B}}O_G$ are known.

$$l_{{}^{\mathrm{B}}O_{\mathrm{G}}\mathrm{N}} = ({}^{\mathrm{B}}x_{\mathrm{N}} - {}^{\mathrm{B}}x_{{}^{\mathrm{B}}O_{\mathrm{G}}}, {}^{\mathrm{B}}y_{\mathrm{N}} - {}^{\mathrm{B}}y_{{}^{\mathrm{B}}O_{\mathrm{G}}}, {}^{\mathrm{B}}z_{\mathrm{N}} - {}^{\mathrm{B}}z_{{}^{\mathrm{B}}O_{\mathrm{G}}}) \tag{15}$$

$$\zeta_i = \arcsin\left(\frac{l_{{}^{\mathrm{B}}O_{\mathrm{G}}\mathrm{N}} \cdot n_{\mathrm{A}_1\mathrm{A}_3{}^{\mathrm{B}}O_{\mathrm{Gh}}}}{\left|l_{{}^{\mathrm{B}}O_{\mathrm{G}}N}\right| \cdot \left|n_{\mathrm{A}_1\mathrm{A}_3{}^{\mathrm{B}}O_{\mathrm{Gh}}}\right|}\right) \tag{16}$$

The angle between $\mathrm{A}_1\mathrm{A}_3$ and the horizontal plane ξ_i can be determined similarly. The degree of stability with $\mathrm{A}_1\mathrm{A}_3$ as the boundary condition can be calculated from the definition of NESM, as follows:

$$E_i = D_i(1 - \cos\zeta_i)\cos\xi_i \tag{17}$$

Based on the definition of NESM, the minimum value of the degree of stability for each adjacent support boundary was considered. Thus, the stability margin is defined as:

$$E = \min_{i}^{n} E_i = D_i(1 - \cos\zeta_i)\cos\xi_i \tag{18}$$

The robot was considered stable when $E > 0$, with a larger E indicating greater stability. Therefore, when optimizing the poses of the robot, a stability margin threshold $E_{\min}$ should be set, and the stability margin at any moment of the motion should be greater than $E_{\min}$.

2.2.3. Adaptive Attitude in Terms of Energy Consumption

The hydraulic driving force of the legs of the hexapod robot was calculated, with Legs 1, 3, and 5 serving as support legs. Based on the leg structure and the planned foot force trajectory, the position of the foot-end relative to the base joint and the magnitudes of the large and small leg drive forces were analyzed for a constant foot-end force; the results for Leg 5 are shown in Figure 5a,b. Because the hexapod robot leg is hydraulically driven, the driving force of the hydraulic system is determined by the maximum load, which is given by the force on the side with the higher force. The relationships between the hydraulic-system drive force F and the geometric parameters H_0 and B_0 are shown in Figure 5c. It was observed that when B_0 is greater than 0.65 m and less than 1.75 m, the hydraulic system driving force F changes uniformly; with an increase in H_0, it initially decreases and subsequently increases, and it gradually increases with an increase in B_0.

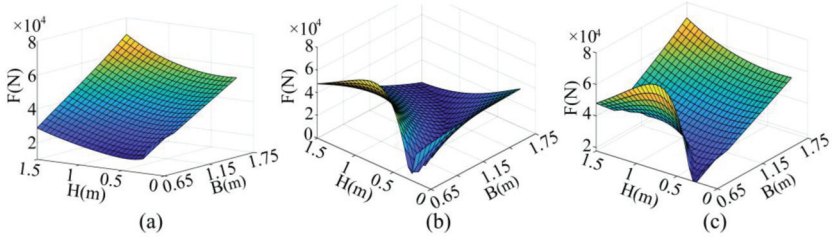


Figure 5. Analysis of hydraulic drive forces for the legs. (**a**) Thigh cylinder. (**b**) Shin cylinder. (**c**) Hydraulic driven force.

In this study, we analyzed the impact of the center-to-ground height H of the hexapod robot torso form, torso-pitch angle α, and roll angle β on the energy consumption during a four-legged support movement (Leg 1, Leg 2, Leg 4, and Leg 5) across different terrains. Three terrain scenarios (φ_x = 5°, φ_y =30°; φ_x = 10°, φ_y =20°; and φ_x = 10°, φ_y =10°) were investigated.

The results shown in Figure 6 indicate that when the pitch angle φ_y of the terrain is small, and the α and β values are close to the terrain slope, the robot consumes less energy during one gait cycle. As the terrain slope decreases, the torso of the robot becomes parallel to the incline, resulting in an increased side-swing cylinder force and a reduced force on the thigh and lower leg cylinders. However, the cylinder pressure remained high at the base junction because the thigh and leg cylinders primarily supported the weight of the robot. The output pressure of the hydraulic system depends on the maximum load, resulting in minimal energy consumption at this stage. As per the analysis in Figure 5, the force required to maintain system balance varies under different postures. When the torso stance was not parallel to the support surface, a difference was observed in C_{zj3} between the endpoints of the feet of the support legs in the torso coordinate system, which exacerbated the uneven output force of the hydraulic cylinders of each support leg and increased the energy consumption of the hydraulic system. In contrast, when the terrain slope angle was large, the maximum output pressure of the hydraulic system depended on the output pressure of the lateral swing cylinders, resulting in an increase in the output pressure of the hydraulic system.

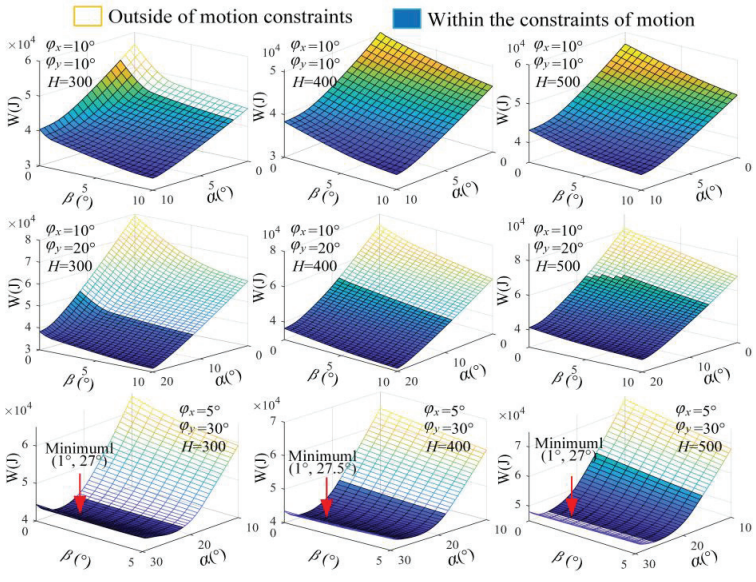


Figure 6. Relationship between posture and energy consumption on different terrains.

As the center-to-ground height of the torso form increased, the energy consumption initially decreased and then increased. This is because the maximum force of the hydraulic system decreased and then increased with the variations in height H. Moreover, as height H increased, the range of available torso-pitch and roll angles also increased, which were constrained by the torso safety zone size and drive hydraulic-cylinder position limits. When the terrain slope increased, the stability constraint became dominant, restricting the torso's attitude and reducing the range of selectable attitudes.

Different levels of energy consumption were observed for different terrains and body postures. On a rough terrain, the individual support legs required unique footing points, and legs with higher posture forces experienced smaller foot-end force distributions to minimize the energy usage of the overall system. Therefore, we optimized the postures, with the objectives of reduced energy consumption and increased use of the torso's kinematics, and selected the appropriate torso postures so that the hydraulic drive force would be as small as possible and the torso's workspace would be as large as possible. The optimization method was to collect multiple sets of simulation data as samples and use them to train a BP neural network model. The input was the terrain angle formed by the supporting feet and the initial height of the body, and the output was the body posture angle with the lowest energy consumption.

3. Force–Position Adaptive Impedance Control of Legs

3.1. Attitude Adaptive Adjustment

Figure 7 presents the overall control diagram of this study, which primarily includes terrain perception, trajectory planning, and controller components. The environment perception component comprises the terrain perception discussed in Section 2.1, as well as the obstacle detection based on the vision module. Following this, corresponding trajectory planning is conducted according to the perceived environment, mainly including safety constraints, stability control, and energy-optimal posture adjustment, as discussed in Section 2.2. Lastly, tracking control is performed on the foot-end force and cylinder position based on the trajectory planning.

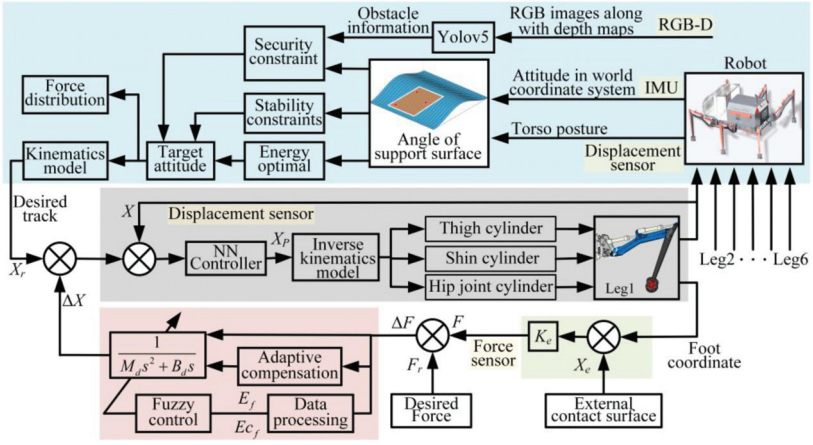


Figure 7. Diagram of motion strategy for heavy-duty hexapod robot.

A force–position composite control was employed to maintain body-posture stability during motion. The control method for the leg is the integration of an outer-loop adaptive fuzzy impedance algorithm with an inner-loop position control. The impedance controller generates the position correction $\Delta \boldsymbol{X}$ based on the difference $\Delta \boldsymbol{F}$ between the contact force F and the desired force Fr, and the motion position X_r is subsequently adjusted with the position correction $\Delta \boldsymbol{X}$. This enables the six-legged robot to maintain its body posture on a complex terrain by tracking the force.

In this study, the inner-loop position control utilized a neural network controller. A neural network controller, which continuously learns through feedback errors, was selected as the position controller; it gradually assumes the dominance of the control behavior. The outputs of the neural network control are as follows:

$$u(t) = u_n(t) + u_f(t) \tag{19}$$

where $u_f(t)$ is the output of the PD control and $u_n(t)$ is the output of the neural network model.

In the early stages of control, the feedback controller plays a predominant role, with the untrained neural network controller continuously learning through feedback errors. This is a process of simultaneous control and learning, possessing the ability to adapt to changes in the object and the environment. In this study, a three-layer feed-forward network has been utilized as the controller. The input–output relationship of the controller can be described as:

$$u_n(t) = W_{OH}^T(t) f\left[W_{HI}^T X_I(t) + \theta_H(t)\right] \tag{20}$$

where $X_I = \{r, e(t), \ldots, e(t-m+1)\}$ is the input vector of the network where r is the reference input; et is the tracking error; $W_{OH}^T(t) W_{HI}^T(t)$ are the weight matrices from the output layer to the hidden layer and the hidden layer to the input layer of the neural network controller; and $\theta_H(t)$ is the input bias of the hidden layer unit. The learning rules of the network weights can be derived according to the feedback error learning method, as:

$$\frac{d[w(t), \theta(t)]^T}{d(t)} = \eta(t) \left| \frac{\partial u_n(t)}{\partial \pi(t)} \right|^T E\left[u_f(t)\right] \tag{21}$$

where $\eta(t)$ is the learning rate; to avoid oscillations in the system, the following adaptive adjustment of the learning rate is used:

$$\eta(t) = 2^{\lambda} \eta(t-1) \tag{22}$$

$$\lambda = sign[\frac{\partial u_n}{\partial[w(t)\ ,\theta(t)]^T} \quad \frac{\partial u_n}{\partial \pi(t-1)}] \tag{23}$$

When the output of the PD feedback controller is smaller than the set threshold, the learning of the neural network ends temporarily, and the neural network controller completely replaces the feedback controller, controlling the system.

3.2. Adaptive Impedance Control

The position-based impedance control model can be conceptualized as a virtual spring model. Mathematically, the impedance control model can be expressed as a function of the robot forces, as follows:

$$\Delta F = M_d \Delta \ddot{X} + B_d \Delta \dot{X} + K_d \Delta X \tag{24}$$

where M_d, B_d, and K_d are the target inertia, target damping, and target stiffness matrices of the impedance control model for the robot, respectively; $\Delta F = F_r - F$, $\Delta X = X - X_r$. F and F_r represent the actual and desired contact forces, respectively, and X and X_r represent the actual and desired positions, respectively. When the foot-end of the robot is in contact with the environment, the impedance outer-loop control generates a position compensation value. This value represents an integrated control object consisting of the terrain environment and the foot-end of the robot. The contact model between the foot-end and the terrain environment can be simplified as follows:

$$F = M_e(\ddot{X} - \ddot{X}_e) + B_e(\dot{X} - \dot{X}_e) + K_{en}(X - X_e) \tag{25}$$

where the stiffness matrix (M_e), damping matrix (B_e), and stiffness matrix of the environment (K_{en}) are fundamental parameters used to describe the interaction of a legged robot with its surrounding terrain. X and X_e indicate the terrain positions.

When the foot-end maintains stable contact with the terrain, the small acceleration and velocity terms in the equation do not significantly affect the foot-end force (F) and can therefore be disregarded.

To simplify the contact model between the foot-end and the terrain environment, this study proposes using a spring model that considers the environmental stiffness (K_e), which is equivalent to the leg stiffness (K_l) and terrain environment stiffness (K_{en}) in series.

$$F = K_e(X - X_e) \tag{26}$$

In this study, we focused on the force in the z-direction, and the following equations were obtained from Equation (25):

$$\begin{cases} \dot{x} = \frac{\dot{f}}{k_e} + \dot{x} \\ \ddot{x} = \frac{\ddot{f}}{k_e} + \ddot{x}_e \end{cases} \tag{27}$$

After substituting Equation (27) into Equation (24), the following equation is obtained:

$$m_d^z \ddot{e}_f + b_d^z \dot{e}_f + (k_d^z + k_e^z) e_f = m_d^z \ddot{q} + b_d^z \dot{q} + k_d^z q \tag{28}$$

where $q = f_r^z + k_e^z x_e - k_e^z x_r$ and owing to the uncertainty of the terrain location, an estimation of environmental position $\hat{x}_e$ is utilized. By setting the desired trajectory $x_r = {}_e$, $q = f_r^z + k_e^z \eta x_e$.

The error function of the force is obtained according to Equation (28):

$$E(s) = \frac{(m_d^z s^2 + b_d^z s + k_d^z) q(s)}{m_d^z s^2 + b_d^z s + k_d^z + k_e^z} \tag{29}$$

Subsequently, the error equation for the steady-state force is:

$$e_{ss} = \lim_{s\to 0} sE(s) = \lim_{s\to 0} \frac{s(m_d^z s^2 + b_d^z s + k_d^z)q(s)}{m_d^z s^2 + b_d^z s + k_d^z + k_e^z} \tag{30}$$

When f_r is a constant force, the steady-state force error is:

$$e_{ss} = \frac{k_d^z(f_r^z + k_e^z \eta x_e)}{k_d^z + k_e^z} \tag{31}$$

Equation (31) shows that to achieve zero steady-state force error in impedance control, either condition $k_d^z = 0$ or $f_r^z = k_e^z \eta x_e$ must be satisfied. However, for a hexapod robot navigating a complex environment, it is challenging to measure the exact equivalent environmental stiffness (k_e) and position (x_e), which makes it difficult to satisfy the condition $f_r^z = k_e^z\ \eta x_e$ in practice. The center-of-gravity of the hexapod robot changed as it walked, resulting in a time-varying force at the foot-end. When f_r is a time-varying force, the steady-state force error still exists, even if $k_d^z = 0$ is set. To mitigate this error, this study proposes an adaptive compensation algorithm to enhance the foot-end tracking performance against time-varying expected forces in the face of environmental uncertainty. Specifically, an adaptive compensation term is introduced into the control scheme.

$$\Gamma(t) = \Gamma(t-k) + \varsigma(f_r^z(t-k) - f^z(t-k)) = \Gamma(t-k) + \varsigma e_f^z(t-k) \tag{32}$$

where ς is the adaptive learning rate and k is the sampling period of the controller.

Assuming $k_d^z = 0$ and incorporating the adaptive compensation term, the impedance model can be reformulated as follows:

$$m_d^z(\ddot{x}(t) - \ddot{x}_r(t)) + b_d^z(\dot{x}(t) - \dot{x}_r(t)) = e_f^z(t) + \Gamma(t) \tag{33}$$

The steady-state force error equation can be re-derived as follows:

$$m_d^z \ddot{e}_f^z(t) + b_d^z \dot{e}_f^z(t) + k_e^z e_f^z(t) + k_e^z \Gamma(t) = m_d^z \ddot{q}(t) + b_d^z \dot{q}(t) \tag{34}$$

Expanding the $\Gamma(t)$, the first term of the expansion is set to zero, after which the expansion is applied to Equations (33) and (34).

$$m_d^z \ddot{e}_f^z(t) + b_d^z \dot{e}_f^z(t) + k_e^z e_f^z(t) + \varsigma k_e^z(e_f^z(t-(n+1)k) + \cdots + e_f^z(t-k)) = m_d^z \ddot{q}(t) + b_d^z \dot{q}(t) \tag{35}$$

This equation can be transformed using Laplace to obtain the transfer function for the steady-state force error.

$$e_f(s) = \frac{m_d^z s^2 + b_d^z s}{m_d^z s^2 + b_d^z s + k_e^z + \varsigma k_e^z(e^{-(n+1)ks} + \cdots + e^{ks})} q(s) \tag{36}$$

The stability condition of the system can be determined using the Rouse stability criterion.

$$0 < \varsigma < \frac{k b_d^z}{m_d^z + k b_d^z} \tag{37}$$

By substituting the constant force expectation and variable expectation force into the steady-state force error function as input signals, the corresponding table for the steady-state force error can be obtained, as shown in Table 2. The proposed adaptive strategy enables force tracking to achieve convergence and eliminate steady-state errors.

Table 2. Steady-state error table for force tracking.

	$f_r = A$	$f_r = Bt^2$	$f_r = ct^2$
IC	$k_d A/(k_e + k_d)$	∞	∞
$K_d{}^Z = 0$	0	$b_d B/k_e$	∞
AD	0	0	$m_d c/\sigma k_e$

Figure 8 depicts the joint simulation system constructed using MATLAB/Simulink and Adams to analyze the influence of m_d, b_d, and k_e on the control performance of the system. Figure 8c shows that as the target impedance inertia parameter m_d decreases, the response speed increases, while the regulation time and overshoot decrease. However, when the overshoot approaches a zero value, m_d decreases again, leading to a decrease in the response speed and an increase in the time required to reach a steady state. Figure 8b indicates that increasing b_d results in a decrease in the response speed and regulation time, whereas the absence of an overshoot leads to a decline in the response speed. However, when an overshoot is absent, increasing b_d leads to a decrease in the response speed. Figure 8a reveals that changes in the environmental stiffness parameters result in significant variations in the force-tracking performance. Therefore, we proposed a real-time self-adaptive fuzzy algorithm to adjust the impedance model parameters based on the error information determined between the expected and actual contact forces. This approach enhances the force-tracking performance of the impedance controller under different environmental parameters, thereby improving the adaptability of a six-legged robot to its surroundings.

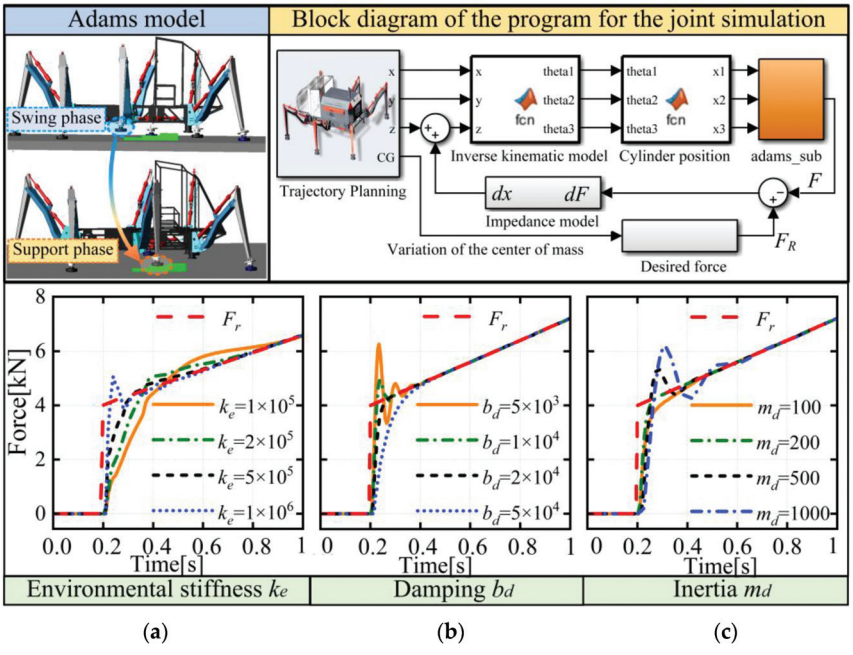


Figure 8. Analysis of the influence of impedance parameters.

The primary challenge in creating an impedance control system involves the selection of appropriate impedance control parameters to improve adaptability to varying environments. Therefore, we proposed a real-time self-adaptive fuzzy algorithm [29] to adjust the impedance model parameters based on the error information determined between the expected and actual contact forces.

The fuzzy control rules for the proposed method are as follows. When the error and error-change rate between the desired force at the foot-end and the contact force are

significant, the target damping parameter should be small to enhance the system response. Conversely, when the error and error change rate between the desired force at the foot-end and contact force are negligible, the target damping parameter should be large to minimize the overshoot of the system and suppress its oscillation. Because the target inertia parameter md is sensitive to the dynamic response, and the effects of md and b_d on the control system are relatively similar, we adjusted the target damping parameter b_d to reduce the complexity of the control system. The input of the fuzzy controller is the error E_f between the expected force f_r at the foot-end and the contact force F, and the rate of change Ec_f of the force error is calculated using Equation (38). The output of the fuzzy controller is an increment in the target damping parameter of the impedance controller.

$$\begin{cases} E_f = (F_r - F) \\ Ec_f = \frac{E_f}{\kappa} \end{cases} \tag{38}$$

where κ represents the sampling period.

The theoretical domain of the input and output quantities of the fuzzy controller is considered as $[-1, 1]$, and the corresponding normalization coefficients transform the ranges of the input and output quantities into the theoretical domain interval. The output of the fuzzy controller is an increment in the target damping parameter of the impedance controller.

$$\begin{cases} \widetilde{E}_f = \frac{2}{E_{f\max} - E_{f\min}} \cdot (E_f - \frac{E_{f\max} + E_{f\min}}{2}) \\ \widetilde{E}c_f = \frac{2}{Ec_{f\max} - Ec_{f\min}} \cdot (Ec_f - \frac{Ec_{f\max} + Ec_{f\min}}{2}) \\ \Delta b = \frac{\Delta\tilde{b}(\Delta b_{\max} - \Delta b_{\min})}{2} + \frac{\Delta b_{\max} + \Delta b_{\min}}{2}) \end{cases} \tag{39}$$

where $E_{f\max}$, $E_{f\min}$, $Ec_{f\max}$, $Ec_{f\min}$, $\Delta b_{\max}$, and $\Delta b_{\min}$ are the maximum and minimum values of E_f, Ec_f and Δb, respectively.

The fuzzy inference process, as shown in Figure 9, involves dividing the various input and output quantities of the fuzzy controller into seven fuzzy levels [NB, NM, NS, ZE, PS, PM, PB], which increase in order from NB to PB. The Gaussian-type subordination function was selected as the subordination function. The fuzzy control rules and image of the affiliation function for the target damping parameter adjustment were established according to the principle of target damping parameter adjustment. The parallel method was used for fuzzy inference, and the center-of-gravity method was used for defuzzification. The increment of the normalized target damping parameter $\Delta\tilde{b}$ was obtained.

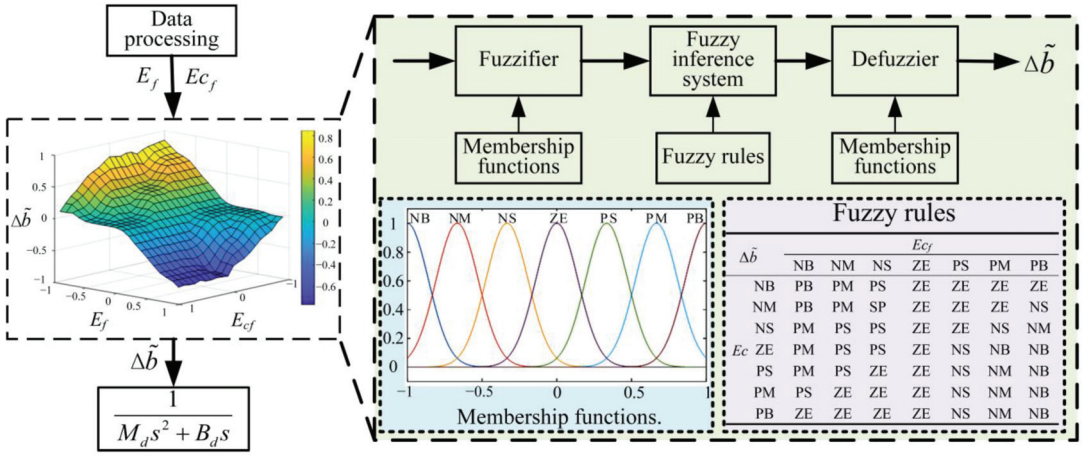


Figure 9. Flowchart of the fuzzy-logic control.

3.3. Force Tracking Experiment

To validate the ability of the force-tracking algorithm to operate on real robots on different terrains, this study conducted experimental investigations on terrain obstacles of varying environmental stiffness. The torso height of the robot was set to 0.4 m above the ground, the gait period was 2 s. The experimental setup involved using the five legs of the hexapod robot as supports while moving the remainder leg within three different terrain environments: hard ground, soft foam ground, and soil. The movements included swing, contact, and foot force tracking. The inertia parameter was set to md = 100 kg, and the input quantity ranges of the fuzzy controllers E_f and Ec_f were [−5000, −5000] and [−1000, 1000], respectively. The initial value of the damping parameter was bd = 2000 N/(m/s), and the output damping parameter adjustment range was [−5000, 5000]. Figure 10 illustrates the experimental results of the robot on different terrains by comparing different force-tracking algorithms.

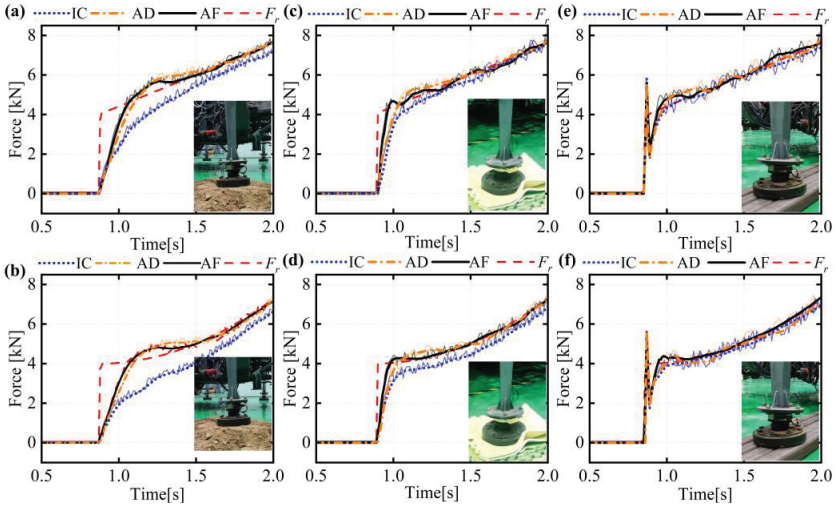


Figure 10. Performance of force tracking for different environmental stiffnesses. (**a**) Slope force tracking in soft soil terrain. (**b**) Slope force tracking in soft foam terrain. (**c**) Slope force tracking in hard terrain. (**d**) Acceleration force tracking in soft soil terrain. (**e**) Acceleration force tracking in soft foam terrain. (**f**) Acceleration force tracking in hard terrain.

The force-tracking curves in different terrain environments displayed different responses, as shown by the experimental results. In hard-terrain environments, where the landing impact was greater, both foot force tracking algorithms exhibited faster responses, as shown in Figure 10e,f. The adaptive compensation algorithm (AD) for force tracking showed a response process similar to that of the traditional impedance control (IC), owing to the large equivalent terrain stiffness of the hard terrain, which resulted in a small steady-state error. On the soft-foam terrain (Figure 10c,d), where the impact force when the foot-end falls to the ground is low, the IC algorithm tracking of the desired acceleration force showed a state of nonconvergence of the force steady-state error. However, as the leg was compressed, the equivalent environmental stiffness gradually increased, and the steady-state error decreased. The AD algorithm improved the response speed and exhibited a relatively small steady-state force error. As shown in Figure 10a,b, the impact force on the foot-end hitting the ground is negligible on the soil. Owing to the lower equivalent stiffness of the soil terrain environment, the response of the IC algorithm becomes slower, and the steady-state force error in force tracking increases. The steady-state error decreases as the soil is compressed, and the equivalent environmental stiffness increases.

The AD algorithm tracks the slope expectation force with a faster response rate and convergence in the steady-state error of force for tracking the acceleration expectation force. The performance of the force-tracking effect is better than that of the IC, which not only improves the response speed, but also results in a relatively small force–steady-state error. However, the adaptive fuzzy (AF) algorithm demonstrated improved performance in force tracking, resulting in a better response speed and reduced force overshoot. Overall, the introduction of the response process increased the amount of overshoot in the system response compared with that without the AD. The introduction of the AF algorithm resulted in relatively better tracking of slope forces at the foot-end of the leg.

4. Experimental Studies and Discussion

The experimental platform utilized for the study of the six-legged robot discussed in this study is depicted in Figure 11. The hexapod had a total mass of 3.55 metric tons, with an electrohydraulic servo mechanism based on hydraulic propulsion. Once the desired trajectory was established, the corresponding current value was input to the servo valve to regulate the movement of the hydraulic cylinder associated with the valve, which was equipped with a built-in displacement sensor to provide feedback regarding its position. A hydraulic pump supplied by Danfoss offered a maximum output pressure of 31 MPa.

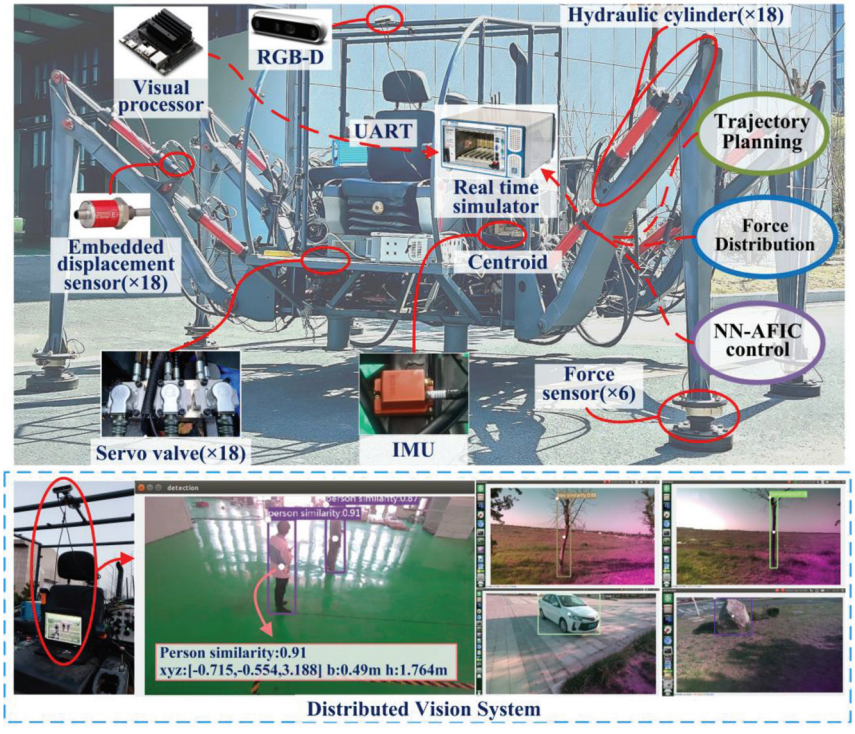


Figure 11. Experimental platform for heavy-duty transport robot.

The robot features a force sensor and an attitude and heading reference system to enable monitoring of the vehicle's status. The distributed visual perception system was set up using the Nvidia Jetson Nano processor and the PyTorch framework. The YOLOv5 algorithm [30] was used as a neural network algorithm. The categories of obstacles were set as tree trunks, people, cars, and boulder. The samples of relevant detection results are shown in Figure 10, and the information of obstacles was transmitted to the motion controller through serial communication. Following calibration, the error between the

obstacle coordinates obtained from the vision module and the actual coordinates was found to be less than 8 mm. This level of precision enables the supply of obstacle information for the obstacle-avoidance strategy outlined in Section 2.2.1. The control approach involves a hardware-in-the-loop system that utilizes an embedded control host powered by the VxWorks operating system with a control cycle of 1 ms and real-time data transmission facilitated through UDP communication to the host computer. The multi-sensor data type and driver configurations were established on the PCI board, and the program's real-time simulation model and driver were developed using MATLAB/Simulink. The simulation program was coded using Microsoft Visual Studio. The structure of the neural model in this experiment was set as $3 \times 8 \times 1$, the initial weights of the neural network were randomly selected within $[-0.1, 0.1]$, the learning rate was set as 0.07, and the value range of the neuron selection sigmoid activation function was $[-1, 1]$. The parameters of the adaptive fuzzy impedance were set to be the same as the experimental parameters in Section 3.3.

First, the effect of the force–position composite control algorithm on the posture of the torso during motion was compared. The experimental scenario is shown in Figure 12. The full cycle consisted of three steps, with two legs for swing in each step and the remaining four legs for support. This gait style is more suitable for complex terrain, because it has higher stability than three-legged support. The base junction foot-end phase began at an initial height of -0.4 m, and the motion step length was 0.8 m. Each step took 3 s, with a 0.3 m lift in the z-direction when swinging. The motion trajectory of the robot used a low-impact foot-end trajectory method that facilitated the initial position and velocity constraints by employing acceleration as a sinusoidal function.

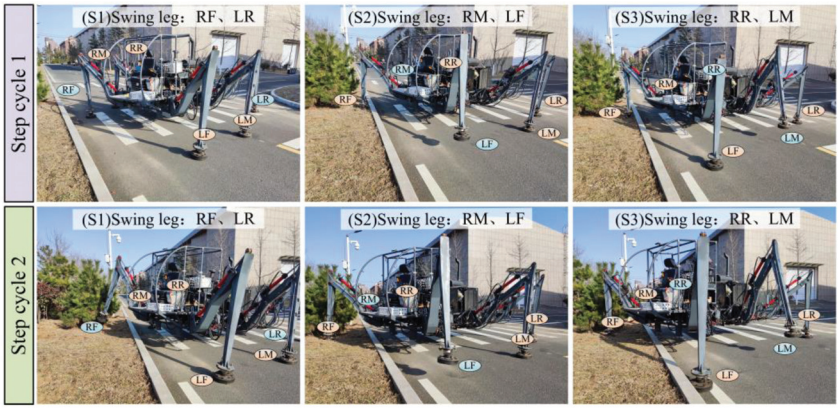


Figure 12. The motion process of the hexapod robot in complex terrain.

The trajectory of the leg on one side of the hexapod robot during position-based control relative to its initial stance in the forward Y-direction and vertical Z-direction is shown in Figure 13a. Each side of the robot has two support-legs and one swing-leg involved in each gait cycle, and the force on each leg is shown in Figure 13b, where the right front (RF) leg and right middle (RM) leg step on the obstacle at 2.73 s and 5.45 s, respectively. The larger force generated when the obstacle is stepped on also affects the forces on the other legs, which has a greater impact on the torso stance of the body. During the support process, owing to the uncertainty of the terrain, the position of the center-of-mass changes accordingly at different slopes, leading to a large deviation in the force of each leg from the expected force, when the posture of the robot torso is influenced by the environment. The robot's torso posture is shown in Figure 13c, and due to the irregular terrain of the landing point, the robot's roll and pitch angles have undergone significant changes. Figure 13d,e show the z-direction coordinates of the legs in the force–position control situation and the forces acting on the legs; it can be seen that when the obstacles are encountered during the swing phase at 2.65 s and 5.68 s, the sensor detects the changes in force and avoids the

obstacles by adjusting the displacement in the z-direction, which is adjusted to avoid the movement of the entire vehicle. In the support phase, the change in the displacement in the z-direction is adjusted according to the change in the support force, thus ensuring that the desired force of the leg to the torso support agrees better, and that the foot-end force can fluctuate around the desired force to achieve dynamic equilibrium. The posture of the robot's trunk is shown in Figure 13f, and it can be seen that the robot can maintain stability of the trunk after a brief adjustment during phase switching. This indicates that the control method proposed in this article can enable the robot to maintain the target attitude while moving in complex terrain.

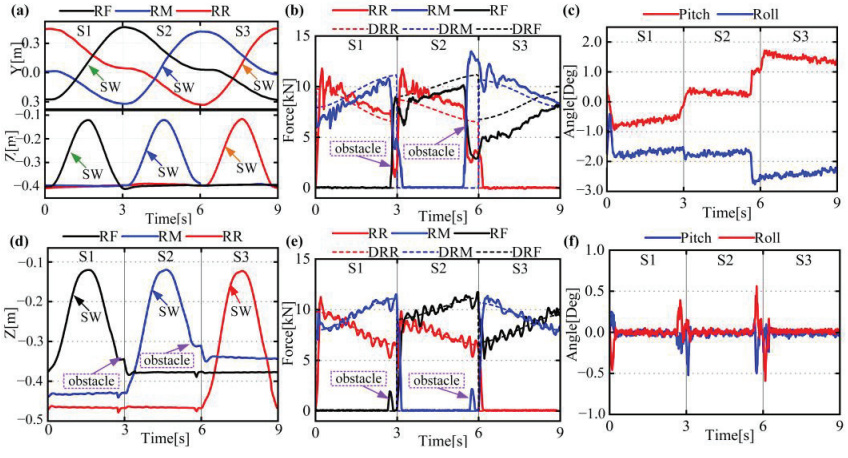


Figure 13. Comparison of motion states for mixed force–position control. (**a**) End-of-foot tracking by position control. (**b**) Force at the end of the foot controlled by position (**c**) Angle of the trunk controlled by position. (**d**) End-of-foot tracking via position-force control. (**e**) End-of-foot force tracking via position-force control. (**f**) Angle of the torso via position-force control.

To validate the posture optimization method adapted for various terrains, experiments were conducted using two types of terrains: the hillside, with a pitch angle of 10°, and asphalt pavements which had a pitch angle of 14°, as shown in Figure 14. Experiments were conducted with the hexapod robot walking on a slope at different trunk center-of-mass heights (H) and trunk pitch angles (φ_y). The robot's total energy consumption for each stride was determined through the pressure and flow of the hydraulic system. The system's working pressure was obtained through a pressure sensor, and the working flow was calculated based on the cylinder displacement. As shown in Figure 15, it can be observed that under two slope conditions, the energy consumption of the hexapod robot with the growth of H exhibits a trend of initially decreasing and then increasing. Moreover, as the trunk pitch angle increases, the energy consumption gradually decreases. At a terrain pitch angle of $\varphi_y = 10°$ and a trunk center-of-mass height of H = 400 mm, the minimum energy consumption occurred when the trunk pitch angle was at 10°. The theoretically calculated result was 34.2 kJ, while the experimental result was 36.1 kJ. When the estimated slope was at a pitch angle of $\varphi_y = 14°$, the minimum energy consumption of the hexapod robot occurred at a trunk center-of-mass height of H = 400 mm and a trunk pitch angle of 14°. The theoretical result was 33.1 kJ, and the experimental result was 35.0 kJ. The actual energy consumption of the robot hydraulic system in each step is shown in Table 3. These results indicate a certain discrepancy between the theoretically calculated energy consumption and the actual values. This discrepancy is due to additional energy consumption caused by pressure shocks and position errors during the actual motion process. However, the trend of energy consumption changes with posture being observed in actual walking is

consistent with the theoretically calculated energy consumption, thereby validating the effectiveness of the posture optimization strategy proposed in this study.



Figure 14. Experimental scenarios in different stiffness environments.

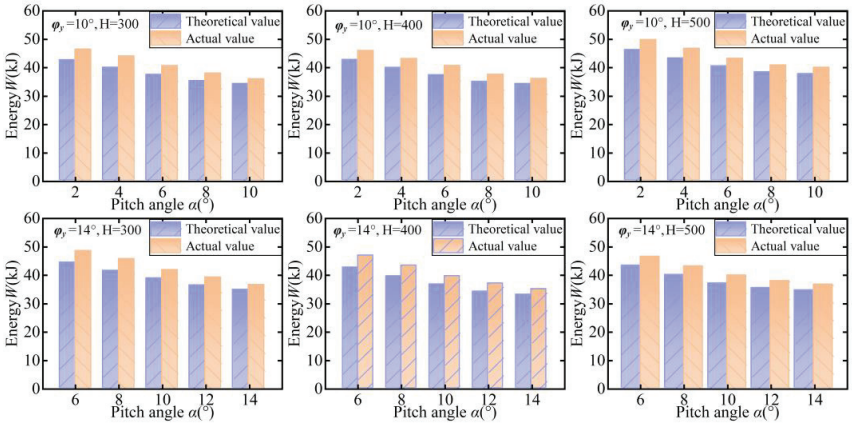


Figure 15. Performance comparison of different force-tracking algorithms.

Table 3. Table of actual energy consumption.

	Pitch Angle	H = 200 mm	H = 300 mm	H = 400 mm	H = 500 mm
φ_y = 10°	0°	-	-	49.7	54.2
	2°	-	46.4	45.9	49.8
	4°	43.1	44.0	43.1	46.7
	6°	40.3	40.6	40.6	43.2
	8°	38.7	37.9	37.5	40.8
	10°	36.3	35.9	36.1	40.0
φ_y = 14°	2°	-	-	-	55.2
	4°	-	-	50.1	50.8
	6°	-	48.6	46.9	46.6
	8°	45.7	45.7	43.4	43.2
	10°	42.7	41.8	39.6	39.9
	12°	39.5	39.2	37.0	38.0
	14°	37.9	36.5	35.0	36.6

To validate the performance of the proposed self-adaptive fuzzy impedance control method for actual robot motion, we conducted a comparative analysis of the proposed method with the impedance control method (ICM) used in the previous reference [15]. As shown in Figure 14, the robot fell into the hillside pavement while moving, indicating that the hillside pavement was less stiff than the rigid asphalt pavement Because terrain

uncertainty has a significant impact on the comparative study results, a transition phase was added in the phase-switching process by introducing a six-legged co-supported gait before the switching phases, and a constant force was set as the expected force of the support leg during this phase.

The experimental results are shown in Figure 16. The right rear (RR) leg of the robot was the focus of this study, and a three-step motion was used. The movements proceeded as follows: 1–2 s of the swing phase, where the expected force is zero; 2–2.5 s of the transition phase, where the RR leg serves as the support for the next phase of the gait, and the expected force is constant at 6000 N; and 2.5–5 s of the support phase, where the expected force is obtained through trajectory planning and foot force distribution algorithms, resulting in a variable force at variable acceleration. As shown in Figure 16a, both control algorithms exhibit faster responses on hard terrain. The steady-state error in the force did not differ significantly. Figure 16b,c shows the displacement compensation of the two control algorithms and the pitch angle of the robot body in the z-direction at this stage. The tilt angle of the robot body is smaller during the switching phase, and the control effects of the two algorithms are approximately the same.

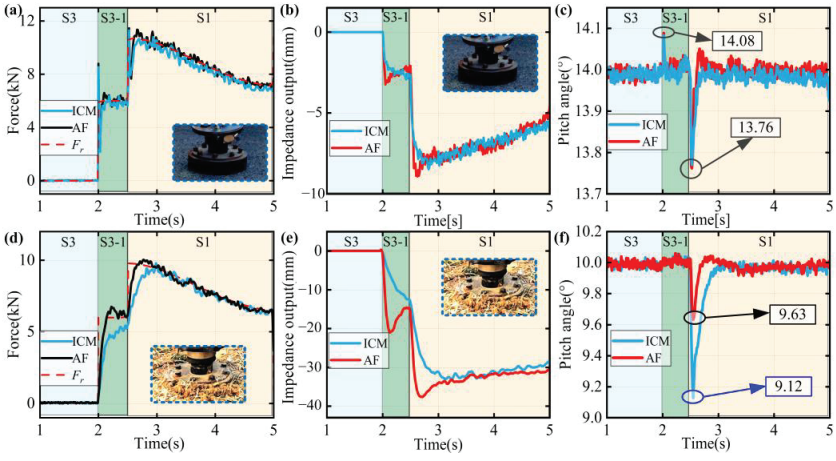


Figure 16. Performance comparison of different force tracking algorithms. (**a**) force tracking in hard terrain, (**b**) position compensation in hard terrain, (**c**) torso angle in hard terrain, (**d**) force tracking in soft terrain, (**e**) position compensation in soft terrain, (**f**) torso angle in soft terrain.

The performance of legged robots is significantly affected on hillside terrain, where the ground stiffness is low. The experimental results depicted in Figure 16d show that, during the transition phase, the ICM is more sensitive to the environment, causing a slower response while tracking the constant force and resulting in force-tracking errors of constant-force tracking during the gait time. In the support phase, the traditional impedance method establishes the target support force in 0.43 s, while the AF control algorithm proposed in this study establishes the target support force in 0.16 s. Both control algorithms exhibit improved desired force-tracking effects, because the robot compressed the soil during the motion, owing to self-weight. This improved the terrain stiffness, particularly by adding a transition phase, thereby reducing the steady-state error. As shown in Figure 16f, in the low-stiffness environments and during phase switching, the maximum tilt angle of the body is 0.37°, whereas the maximum tilt angle of the impedance control is 0.88°. Because of the small compensation amount of the ICM during phase switching, the body experienced a brief tilt instability caused by the low stiffness of the ground, thereby failing to establish support in time.

The rugged terrain and asphalt pavement are considered as soft and hard terrain, respectively, and the experimental results are shown in Table 4. The results demonstrated

that the AF algorithm exhibits improved performance in force tracking when the terrain environment is relatively "soft", whereas the performance improvement of the AF algorithm is less significant when the terrain environment is "hard". Moreover, it was observed that the convergence process of the ICM for force tracking becomes slower as the terrain environment becomes relatively "soft", whereas the AF algorithm demonstrates better performance in improving force tracking in such environments and exhibits certain adaptability to the terrain. However, compared with the single-leg study, the overall tracking performance of the robot experiment decreased and fluctuated owing to the influences of the position-control accuracy and rigid-flexible coupling characteristics of the heavy-duty robot itself, as well as force-tracking errors caused by uncertainties.

Table 4. Table of force-tracking performance.

		Maximum Error of Pitch Angle	Steady-State Time
Hard terrain	ICM	0.25°	0.11 s
	AF	0.24°	0.15 s
Soft terrain	ICM	0.88°	0.43 s
	AF	0.37°	0.16 s

5. Conclusions

In this paper, we propose a motion strategy for agricultural transportation robots operating on complex terrains. The premise of the motion strategy looks to torso safety and stability constraints based on the terrain estimation of the support surface, and a visual recognition system to ensure the safe passage of the robot on different terrains. Moreover, force analysis of the hydraulic leg drive was conducted to establish the correspondence between the driving force and the motion posture. It can be concluded that adjusting the robot's posture can reduce the energy consumption of robot movements. A control method combining force and position was subsequently proposed, maintaining the target posture of the robot's torso through force tracking. To address the problem of tracking time-varying forces in uncertain environments, an adaptive-fuzzy impedance algorithm was developed to eliminate steady-state errors. The proposed control strategy was validated by building a real-time simulation system for a heavy-duty six-legged robot and comparing it with an existing model. Experimental and simulation results showed that the robot maintained the stability of its locomotor posture by tracking the forces on each leg, and the proposed control strategy achieved convergence and shorter response times on various terrains. The proposed control method is able to reduce the adjustment time of the system by 58% and the deflection error of the structure by 62.8% when the stiffness of the external environment is low.

Several issues remain to be resolved before the advent of the widespread use of multi-legged hydraulic robots. These include the fact that hydraulic drive units in swing and support states require different driving forces, but the solution provided by a single-pump hydraulic systems fails to satisfy the differing requirements of multiple legs, leading to energy loss. Another issue is that the flexible deformation of the legs in large-sized, heavy-load robots makes force tracking difficult for the system. Therefore, in the near future, we plan to modify the structure of our hydraulic system and optimize our control algorithms to overcome these challenges.

Author Contributions: Conceptualization, X.L.; investigation, X.L. and C.L.; writing—original draft preparation, K.Y.; writing—review and editing, K.Y. and C.L.; formal analysis, K.Y. and Z.W. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by The National Key Research and Development Program of China, (Grant No. 2016YFC0802904), National Natural Science Foundation of China (Grant No. 52205565), and the China Postdoctoral Science Foundation (No. 2023M730784).

Institutional Review Board Statement: Not applicable.

Data Availability Statement: No new data were created or analyzed in this study. Data sharing is not applicable to this article.

Acknowledgments: We would like to sincerely thank all of the authors who submitted papers to the Special Issue of *Agriculture* entitled "Robots and Autonomous Machines for Agriculture Production", the reviewers of these papers for their constructive comments and thoughtful suggestions, and the editorial staff of *Agriculture*.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Yuan, J.; Ji, W.; Feng, Q. Robots and Autonomous Machines for Sustainable Agriculture Production. *Agriculture* **2023**, *13*, 1340. [CrossRef]
2. Tiozzo Fasiolo, D.; Scalera, L.; Maset, E.; Gasparetto, A. Towards Autonomous Mapping in Agriculture: A Review of Supportive Technologies for Ground Robotics. *Robot. Auton. Syst.* **2023**, *169*, 104514. [CrossRef]
3. Gai, J.; Xiang, L.; Tang, L. Using a Depth Camera for Crop Row Detection and Mapping for Under-Canopy Navigation of Agricultural Robotic Vehicle. *Comput. Electron. Agric.* **2021**, *188*, 106301. [CrossRef]
4. Vrochidou, E.; Tsakalidou, V.N.; Kalathas, I.; Gkrimpizis, T.; Pachidis, T.; Kaburlasos, V.G. An Overview of End Effectors in Agricultural Robotic Harvesting Systems. *Agriculture* **2022**, *12*, 1240. [CrossRef]
5. Ferreira, J.; Moreira, A.P.; Silva, M.; Santos, F. A Survey on Localization, Mapping, and Trajectory Planning for Quadruped Robots in Vineyards. In Proceedings of the 2022 IEEE International Conference on Autonomous Robot Systems and Competitions (ICARSC), Santa Maria da Feira, Portugal, 29–30 April 2022.
6. Song, X.; Zhang, X.; Meng, X.; Chen, C.; Huang, D. Gait Optimization of Step Climbing for a Hexapod Robot. *J. Field Robot.* **2021**, *39*, 55–68. [CrossRef]
7. Lee, W.-J.; Orin, D.E. The Kinematics of Motion Planning for Multilegged Vehicles over Uneven Terrain. *IEEE J. Robot. Autom.* **1988**, *4*, 204–212. [CrossRef]
8. Xu, K.; Lu, Y.; Shi, L.; Li, J.; Wang, S.; Lei, T. Whole-Body Stability Control with High Contact Redundancy for Wheel-Legged Hexapod Robot Driving over Rough Terrain. *Mech. Mach. Theory* **2023**, *181*, 105199. [CrossRef]
9. Choi, S.; Ji, G.; Park, J.; Kim, H.; Mun, J.; Lee, J.H.; Hwangbo, J. Learning Quadrupedal Locomotion on Deformable Terrain. *Sci. Robot.* **2023**, *8*, eade2256. [CrossRef] [PubMed]
10. Kalakrishnan, M.; Buchli, J.; Pastor, P.; Mistry, M.; Schaal, S. Learning, Planning, and Control for Quadruped Locomotion over Challenging Terrain. *Int. J. Robot. Res.* **2010**, *30*, 236–258. [CrossRef]
11. Grzelczyk, D.; Stanczyk, B.; Awrejcewicz, J. Kinematics, Dynamics and Power Consumption Analysis of the Hexapod Robot During Walking with Tripod Gait. *Int. J. Struct. Stab. Dyn.* **2017**, *17*, 1740010. [CrossRef]
12. Ott, C.; Mukherjee, R.; Nakamura, Y. Unified Impedance and Admittance Control. In Proceedings of the 2010 IEEE International Conference on Robotics and Automation, Anchorage, AK, USA, 3–7 May 2010.
13. Zhu, Q.; Huang, D.; Yu, B.; Ba, K.; Kong, X.; Wang, S. An Improved Method Combined SMC and MLESO for Impedance Control of Legged Robots' Electro-Hydraulic Servo System. *ISA Trans.* **2022**, *130*, 598–609. [CrossRef]
14. Zhu, Q.; Zhang, J.; Li, X.; Zong, H.; Yu, B.; Ba, K.; Kong, X. An Adaptive Composite Control for a Hydraulic Actuator Impedance System of Legged Robots. *Mechatronics* **2023**, *91*, 102951. [CrossRef]
15. Wang, S.; Chen, Z.; Li, J.; Wang, J.; Li, J.; Zhao, J. Flexible Motion Framework of the Six Wheel-Legged Robot: Experimental Results. *IEEE/ASME Trans. Mechatron.* **2022**, *27*, 2246–2257. [CrossRef]
16. Kronander, K.; Billard, A. Stability Considerations for Variable Impedance Control. *IEEE Trans. Robot.* **2016**, *32*, 1298–1305. [CrossRef]
17. Roveda, L.; Iannacci, N.; Vicentini, F.; Pedrocchi, N.; Braghin, F.; Tosatti, L.M. Optimal Impedance Force-Tracking Control Design With Impact Formulation for Interaction Tasks. *IEEE Robot. Autom. Lett.* **2016**, *1*, 130–136. [CrossRef]
18. Deng, Z.; Jin, H.; Hu, Y.; He, Y.; Zhang, P.; Tian, W.; Zhang, J. Fuzzy Force Control and State Detection in Vertebral Lamina Milling. *Mechatronics* **2016**, *35*, 1–10. [CrossRef]
19. Yang, J.; Sun, T.; Yang, H. Spatial Hybrid Adaptive Impedance Learning Control for Robots in Repetitive Interactive Tasks. *ISA Trans.* **2023**, *138*, 151–159. [CrossRef]
20. Wang, W.; Li, Q.; Lu, C.; Gu, J.; Li, A.; Li, Y.; Huo, Q.; Chu, H.; Zhu, M. Impedance Estimation for Robot Contact with Uncalibrated Environments. *Mech. Syst. Signal Process.* **2021**, *159*, 107819. [CrossRef]
21. Huang, B.; Ye, Z.; Li, Z.; Yuan, W.; Yang, C. Admittance Control of a Robotic Exoskeleton for Physical Human Robot Interaction. In Proceedings of the 2017 2nd International Conference on Advanced Robotics and Mechatronics (ICARMT), Ai'an, China, 27–31 August 2017.
22. Yang, C.; Zeng, C.; Fang, C.; He, W.; Li, Z. A DMPs-Based Framework for Robot Learning and Generalization of Humanlike Variable Impedance Skills. *IEEE/ASME Trans. Mechatron.* **2018**, *23*, 1193–1203. [CrossRef]
23. Zeng, C.; Yang, C.; Cheng, H.; Li, Y.; Dai, S.-L. Simultaneously Encoding Movement and sEMG-Based Stiffness for Robotic Skill Learning. *IEEE Trans. Ind. Inform.* **2021**, *17*, 1244–1252. [CrossRef]

24. Zhang, T.; Liang, X.; Zou, Y. Robot Peg-in-Hole Assembly Based on Contact Force Estimation Compensated by Convolutional Neural Network. *Control. Eng. Pract.* **2022**, *120*, 105012. [CrossRef]
25. Modares, H.; Ranatunga, I.; Lewis, F.L.; Popa, D.O. Optimized Assistive Human–Robot Interaction Using Reinforcement Learning. *IEEE Trans. Cybern.* **2016**, *46*, 655–667. [CrossRef]
26. Yang, Y.; Wu, X.; Song, B.; Li, Z. Whole-Body Fuzzy Based Impedance Control of a Humanoid Wheeled Robot. *IEEE Robot. Autom. Lett.* **2022**, *7*, 4909–4916. [CrossRef]
27. Homchanthanakul, J.; Manoonpong, P. Continuous Online Adaptation of Bioinspired Adaptive Neuroendocrine Control for Autonomous Walking Robots. *IEEE Trans. Neural Netw. Learn. Syst.* **2022**, *33*, 1833–1845. [CrossRef]
28. Li, Z.; Ge, Q.; Ye, W.; Yuan, P. Dynamic Balance Optimization and Control of Quadruped Robot Systems With Flexible Joints. *IEEE Trans. Syst. Man Cybern. Syst.* **2016**, *46*, 1338–1351. [CrossRef]
29. Ferraz, R.S.C.; Mello, T.P.; Borges Filho, M.N.; Borges, R.F.O.; Magalhães Filho, S.C.; Scheid, C.M.; Meleiro, L.A.C.; Calçada, L.A. An Experimental and Theoretical Approach on Real-Time Control and Monitoring of the Apparent Viscosity by Fuzzy-Based Control. *J. Pet. Sci. Eng.* **2022**, *217*, 110896. [CrossRef]
30. Lu, A.; Ma, L.; Cui, H.; Liu, J.; Ma, Q. Instance Segmentation of Lotus Pods and Stalks in Unstructured Planting Environment Based on Improved YOLOv5. *Agriculture* **2023**, *13*, 1568. [CrossRef]

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.

agriculture

MDPI

Article

Drivable Agricultural Road Region Detection Based on Pixel-Level Segmentation with Contextual Representation Augmentation

Yefeng Sun [1], Liang Gong [1,2,*], Wei Zhang [1], Bishu Gao [1], Yanming Li [1] and Chengliang Liu [1,2]

[1] School of Mechanical Engineering, Shanghai Jiao Tong University, Shanghai 200240, China; ysun3019@sjtu.edu.cn (Y.S.); zhang_wei@sjtu.edu.cn (W.Z.); gaobishu2021@sjtu.edu.cn (B.G.); ymli@sjtu.edu.cn (Y.L.); chlliu@sjtu.edu.cn (C.L.)

[2] MoE Key Laboratory of Artificial Intelligence, AI Institute, Shanghai Jiao Tong University, Shanghai 200240, China

* Correspondence: gongliang_mi@sjtu.edu.cn

Abstract: Drivable area detection is crucial for the autonomous navigation of agricultural robots. However, semi-structured agricultural roads are generally not marked with lanes and their boundaries are ambiguous, which impedes the accurate segmentation of drivable areas and consequently paralyzes the robots. This paper proposes a deep learning network model for realizing high-resolution segmentation of agricultural roads by leveraging contextual representations to augment road objectness. The backbone adopts HRNet to extract high-resolution road features in parallel at multiple scales. To strengthen the relationship between pixels and corresponding object regions, we use object-contextual representations (OCR) to augment the feature representations of pixels. Finally, a differentiable binarization (DB) decision head is used to perform threshold-adaptive segmentation for road boundaries. To quantify the performance of our method, we used an agricultural semi-structured road dataset and conducted experiments. The experimental results show that the *mIoU* reaches 97.85%, and the *Boundary IoU* achieves 90.88%. Both the segmentation accuracy and the boundary quality outperform the existing methods, which shows the tailored segmentation networks with contextual representations are beneficial to improving the detection accuracy of the semi-structured drivable areas in agricultural scene.

Keywords: semi-structured road detection; contextual representation; pixel-level segmentation; agricultural robot

Citation: Sun, Y.; Gong, L.; Zhang, W.; Gao, B.; Li, Y.; Liu, C. Drivable Agricultural Road Region Detection Based on Pixel-Level Segmentation with Contextual Representation Augmentation. *Agriculture* **2023**, *13*, 1736. https://doi.org/10.3390/agriculture13091736

Academic Editor: José Lima

Received: 12 July 2023
Revised: 24 August 2023
Accepted: 31 August 2023
Published: 1 September 2023



Copyright: © 2023 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (https://creativecommons.org/licenses/by/4.0/).

1. Introduction

Accurate detection of semi-structured roads and unstructured roads has a wide range of applications. Taking agriculture as an example, the agricultural environment is harsh, and there are various safety hazards [1]. With the ageing of the population and the reduction in the labor force, automation and intelligence will be the future development direction of agricultural equipment [2]. Drivable area detection using machine vision is crucial for the autonomous navigation of agricultural robots in natural environments. The road environment has three categories, structured roads, semi-structured roads and unstructured roads [3]. For structured roads, such as well-marked highways, lane detection is commonly used to find drivable areas. However, roads in the agricultural scene are mostly semi-structured with ambiguous boundaries and no artificial markings [4]. The road features are complex and easily disturbed by the environment. Compared with structured road detection, semi-structured road detection in agricultural scenes is more challenging [5].

Semi-structured road detection in complex environments has always been a research hotspot in autonomous navigation. The research on vision-based semi-structured road detection can be divided into feature-based road detection, model-based road detection, and image segmentation-based road detection.

Feature-based road detection methods mainly use salient visual characteristics to detect road regions, such as the vanishing point (VP), color, texture, boundary, and other features. Yang G. et al. [5] proposed a contourlet transform framework for VP detection in a single image for unstructured roads. Shi J. et al. [6] used Gabor filters and particle filters to estimate the road texture direction, and designed a noise-insensitive observation model to vote for vanishing points. Liu Y.B. et al. [7] proposed an unstructured road VP detection solution combining a convolutional neural network and heatmap regression. Hernandez D.E. et al. [8] designed a vision-based road boundary tracking system, which detects and tracks the boundary of unstructured roads based on color differences and texture lines. Liu G. et al. [9] estimated the probability distribution of lane parameters using multiple kernel densities, and presented a partitioned particle filter approach for lane detection and tracking. The advantage of feature-based methods is that they rely less on prior knowledge and are not sensitive to road shapes. The disadvantage is that the requirements for road characteristics are high, and misjudgment is prone to occur when there is environmental interference (e.g., water stains, cracks, and shadows).

Model-based methods construct parametric road boundary models with straight lines or curves, and use road model matching to solve the parameters. Perng J.W. et al. [10] adopted a hyperbolic model to fit the feature points of the lane lines and used a particle filter to track the lanes. Based on the third-order B-spline curve model, Cao J. et al. [11] used the random sample consensus algorithm to fit the lanes and evaluate the fitted curve. Wang K. et al. [12] used Bezier splines to construct a variable road template whose parameters could be solved by the improved RANSAC algorithm. Yuan Y. et al. [13] used a structural support vector machine to discriminate road boundary and non-boundary instances for fitting complete boundary lines. Wang X. et al. [14] proposed an unstructured road detection method based on contour selection, which uses Hough transform and contour detection function to detect lines and edges, and takes the line with the best coincidence as the edge contour. These feature-based methods are generally applied to roads with relatively regular shapes. The advantage is that the performance is robust, and the detected road area is relatively complete. The disadvantage is also apparent. It is difficult to accurately match roads with complex shapes, and the wrong road model selection will lead to detection failure [15].

Segmentation-based methods describe road detection as a classification problem of road regions and background regions. Different segmentation algorithms are used to separate different classes of regions with salient features. Traditional road semantic segmentation algorithms are based on handcrafted features (e.g., SIFT, SURF, HoG, etc.) and classical classifiers (e.g., multiple nearest neighbors [16], random forest classifiers [4], conditional random fields [17], etc.) [18]. Alam A. et al. [16] proposed a road detection system for classifying unstructured roads into road and non-road regions using a multiple nearest neighbor (NN) classifier and a soft voting aggregation approach. Xiao L. et al. [4] proposed a structural random forest-based road detection algorithm that exploits the context of image patches and the structural information of labels to obtain consistent segmentation results. Wang Q. et al. [19] combined depth cues with traditional RGB colors to predict road and non-road regions using context-aware label transfer. Geng L. et al. [20] used the Markov random field to optimize the road detection results according to the relationship between super-pixel neighborhoods. These traditional methods explicitly clustered or grouped the pixels while classifying the patch category. In recent years, semantic segmentation algorithms have witnessed tremendous progress with the development of deep learning. Many deep learning-based semantic segmentation networks have been proposed, such as U-net [21], SegNet [22] and DeepLab series [23–26]. These methods have a wide range of applications and improvements in road segmentation. Eff-UNet [18] used Efficient Net as the encoder feature extractor and decoder of UNet, combining high-level features and low-level spatial information to achieve accurate segmentation. Lane-DeepLab [27] added the ASPP to optimize the encoder–decoder structure, and employed the Semantic Embedding Branch and the Single Stage Headless modules to obtain multi-level semantic features.

The segmentation method based on deep learning has a strong learning ability and good adaptability to the environment, and shows unique advantages in road detection tasks. However, a large number of training samples are required in the early stage. Unfortunately, there are few public datasets of semi-structured roads, which brings difficulties to the model's training.

The semi-structured roads, such as roads and fixed paths in greenhouses, mainly have the following characteristics: (1) there are neither standardized lane lines nor identifiable manual signs in the drivable area. (2) The environment is complex and changeable; water stains, shadows, and changing lighting will interfere with road features. (3) Road boundaries are difficult to distinguish because soil and water stains may blur road boundaries. This paper proposes a high-resolution detection network for semi-structured roads to address the challenge of semi-structured drivable area detection in agriculture. Most encoder–decoder architectures concatenate high-to-low encoders to extract road features gradually, then recover high-resolution from low-resolution feature representations. However, as the network layer deepens, the size of the feature map gradually decreases. This operation inevitably produces blurred feature maps after multiple convolutions and loses some crucial details, which is fatal for semi-structured roads in greenhouse boundary extraction. The backbone of our method adopts HRNet [28] to extract high-resolution road features in parallel at multiple scales. HRNet changes the connection between high-resolution and low-resolution resolutions from series to parallel, thus the high-resolution could representationed throughout the whole network structure. To strengthen the relationship between pixels and their corresponding object regions, we use OCR [29] to enhance the feature representations of pixels. Finally, a DB module [30] is used as a decision head to perform adaptive threshold segmentation of road boundaries. The loss function includes two parts: segmentation loss and threshold map loss. To quantify the performance of the proposed method, we make a dataset of semi-structured roads in agricultural scenes and conduct experiments in agricultural scenes.

To summarize, the contributions of this study are as follows:

(1) This paper proposes a high-resolution detection network for semi-structured roads in facility agriculture. The network uses HRNet to extract high-resolution road features in parallel at multiple scales, then uses OCR to enhance the feature representations, and finally uses a DB decision head to segment road boundaries adaptively.
(2) A loss function is designed, including segmentation and threshold map losses.
(3) A dataset of semi-structured agricultural roads for greenhouses is produced, and the method in this paper is validated.

The rest of the paper is organized as follows: Section 2 presents the high-resolution detection method for semi-structured roads in detail. Then, Section 3 discusses the experimental results. Finally, conclusions and future work are presented in Section 4.

2. Theory and Method

2.1. Architecture of Agricultural Semi-Structured Road Detection Network

The agricultural semi-structured road detection network consists of three main components: high-resolution feature extraction, road-contextual representations, and threshold-adaptive boundary segmentation. To extract high-resolution features of semi-structured roads, we must adopt the HRNet as the backbone to preserve high-resolution features during downsampling. To distinguish the features of drivable and non-drivable regions more clearly, we can use contextual representations to augment road objectness. Finally, an adaptive threshold-learned DB decision head is used to segment road boundaries. The pipeline of the semi-structured road detection network is shown in Figure 1.

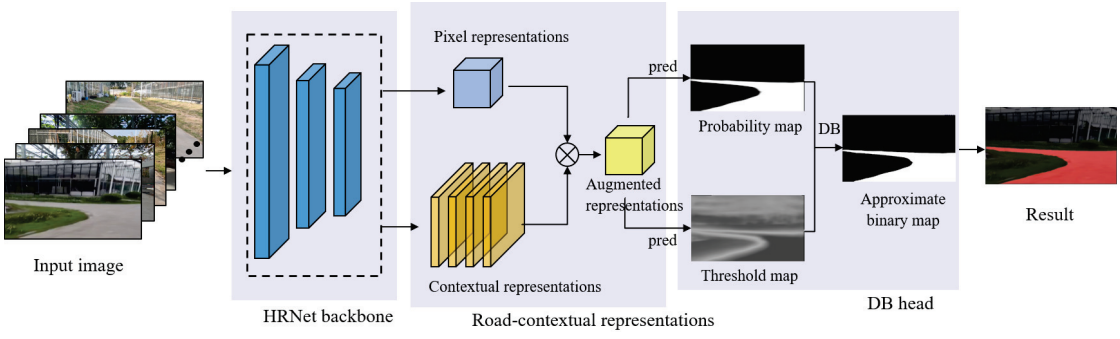


Figure 1. The structure of the agricultural semi-structured road detection network.

2.2. High-Resolution Road Feature Extraction

As mentioned above, the encoder–decoder architectures lead to the insufficient resolution of feature maps after multiple convolutions, thus losing some important shape and boundary details, which is fatal for semi-structured road detection. In contrast, HRNet connects high-to-low resolution convolution streams in parallel rather than in series, which maintains high-resolution representations throughout the whole process.

Therefore, high-resolution feature extraction uses HRNet-W48 as the backbone to extract the semi-structured road features of different levels. The backbone contains four stages with four parallel convolution streams. The first stage includes 4 residual units, each containing a 3×3 convolution and two 1×1 convolutions. The 1×1 convolution is mainly used to change the dimension of feature maps, and the 3×3 convolution is used to extract features. The second, third, and fourth stages include 1, 4, and 3 modularized blocks, respectively. Each branch of the modularized block contains 4 residual units. The modularized block performs upsampling of low-resolution features, downsampling of high-resolution features, and fusion of feature maps at the same level. The multi-resolution fusion of feature maps at different levels can obtain rich road features. For an image, by adding a downsampling branch from high resolution to low resolution, the corresponding number of channels will be doubled, and the resolution will be half. In the fourth stage, the channels of the convolutions of the four resolutions are 48, 96, 192, and 384, and the resolutions are 1/4, 1/8, 1/16, and 1/32. In the final feature fusion stage, we adopt HRNetV2 as the representation head, which uses bilinear upsampling to merge low-resolution and high-resolution representations to obtain pixel representations of u semi-structured roads.

2.3. Road-Contextual Representations

The agricultural environment is complex and changeable. To distinguish the features of drivable and non-drivable regions more clearly, we must enhance the pixel representations with the contextual relationship between pixels and corresponding object regions [28]. The network details are shown in Figure 2. The backbone outputs two parts: pixel representations and soft object regions. The pixel representations are the feature representations of semi-structured roads outputted by the HRNetV2 representation head. In this section, we define two classes: road and non-road. The soft object regions are regions of road and non-road roughly segmented from the backbone through supervised training. Each position of each layer represents the probability that the corresponding pixel belongs to the class represented by this layer. During training, we learn the object region generator under the supervision of the ground-truth segmentation using the binary cross-entropy loss.

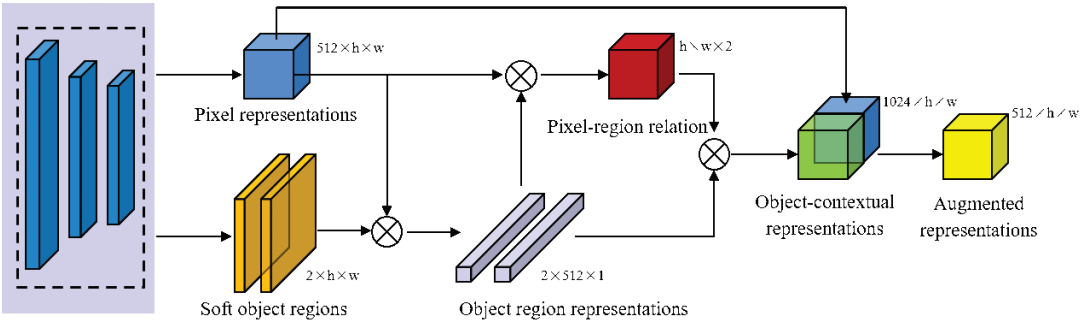


Figure 2. Pixel representations enhanced by object-contextual features.

Object region representations are defined as the aggregation of the feature representations of pixels weighted by their degrees belonging to the k-th object region:

$$\mathrm{f}_k = \sum_{i \in I} \widetilde{m}_{ki} \mathrm{x}_i$$

where f_k is the k-th object region representation and x_i is the representation of pixel p_i of image I. $\widetilde{m}_{ki}$ is the normalized degree for pixel p_i belonging to the k-th object region.

We use self-attention [31] to combine contextual information and pixel representations to obtain object-contextual representations:

$$\mathrm{y}_i = \rho\left(\sum_{k=1}^{2} w_{ik}\delta(\mathrm{f}_k)\right)$$

where y_i is the object-contextual representation of pixel p_i, $w_{ik} = \frac{e^{\kappa(\mathrm{x}_i,\mathrm{f}_k)}}{\sum_{j=1}^{2} e^{\kappa(\mathrm{x}_i,\mathrm{f}_j)}}$ is the relationship between pixel and object regions, $\kappa(\mathrm{x},\mathrm{f}) = \phi(\mathrm{x})^T\psi(\mathrm{f})$ is the unnormalized relation function and $\phi(\cdot)$, $\psi(\cdot)$, $\rho(\cdot)$ and $\delta(\cdot)$ are transformation functions implemented by 1×1 conv $\rightarrow$ BN $\rightarrow$ ReLU.

Augmented representations are a fusion of the object-contextual representations and pixel representations, as demonstrated below.

$$\mathrm{z}_i = g\left(\begin{bmatrix}\mathrm{x}_i^T & \mathrm{y}_i^T\end{bmatrix}^T\right)$$

where z_i is the augmented representation of pixel p_i and $g(\cdot)$ is a transform function implemented by 1×1 conv $\rightarrow$ BN$\rightarrow$ ReLU.

2.4. Threshold-Adaptive Boundary Segmentation

Threshold-adaptive segmentation models road and non-road regions in two channels, calculating and comparing the scores of each pixel belonging to the different regions. The pavement characteristics of semi-structured roads are inconsistent, and the boundaries are not obvious, so it is difficult to set boundary thresholds to segment the road regions. Traditional segmentation methods set a fixed threshold to convert the probability map generated by the segmentation network into a binary image. However, this method is not differentiable and cannot be optimized in the network. Inspired by DBNet++ [30] in scene text detection, we adopt the DB module with an adaptive boundary threshold. Combining the DB module to optimize the segmentation network can obtain highly robust segmentation results.

We perform supervised training on the augmented representations in Section 2.2 to predict the probability map and the threshold map. Combining the probability map and the

threshold map, we use an approximate step function to perform differentiable binarization on each pixel to obtain an approximate binary map.

$$B_{i,j} = \frac{1}{1 + e^{-50(P_{i,j} - T_{i,j})}}$$

where B is the approximate binary map, T is the adaptive threshold map, and P is the adaptive threshold map learned from the network. The function quickly changes from 0 to 1 when the probability is close to the threshold, and this process is derivable. Therefore, the network can adaptively predict the segmentation threshold of each pixel of the road boundary in the image, making the semi-structured road boundary detection more robust.

2.5. Loss Function

This paper has three road segmentation outputs: the soft object region in the backbone, the probability map and the approximate binary map in the DB Head. All three parts can generate road segmentation results, but the fineness of segmentation is different. The classification of road and non-road is a binary classification problem. The segmentation loss of each output applies the binary cross-entropy (BCE) loss, which is defined as

$$L_s = L_p = L_b = -\frac{1}{N}\sum_{i=1}^{N} y_i log x_i + (1 - y_i)\log(1 - x_i)$$

where L_s is the segmentation loss for the soft object region, L_p is the segmentation loss for the probability map, and L_b is the segmentation loss for the approximate binary map, N is the number of pixels, $y_i \in \{0, 1\}$ is the label of the pixel, and $x_i \in [0, 1]$ is the probability of the pixel.

The threshold map of the DB Head can be obtained by learning or supervised training. The value of the threshold map is between 0 and 1, and the network needs to regress to continuous values. We use the L1 loss function as the threshold map loss function, denoted as L_t

$$L_t = -\frac{1}{N}\sum_{i=1}^{N} |y_i - x_i|$$

The complete loss function L consists of two parts: the segmentation loss and the threshold map loss.

$$L = \omega_s L_s + \omega_p L_p + \omega_b L_b + \omega_t L_t$$

where ω_i is the weight parameter of different parts, set to 1 in this paper. The threshold map can be supervised or unsupervised. If the threshold map is not supervised, then ω_t is 0.

3. Experiments

3.1. Dataset

The semi-structured facility agriculture scenario is complex and changeable without the support of open-source datasets. We need to make an agricultural semi-structured road dataset, because network training requires a lot of labeled data. We collected 560 semi-structured road images of agricultural scenes using ZED2 in the glass greenhouses from Shanghai Jiao Tong University and Langxia Town, Jinshan District, Shanghai. We use EISeg [32] to annotate road and non-road regions of the image interactively. To increase the diversity of the samples, we augment the dataset of semi-structured roads with the Road Traversing Knowledge (RTK) dataset [33]. The RTK dataset contains data for different road conditions, and we selected 700 images containing semi-structured roads. The 12 semantic categories of the RTK dataset are mapped into two categories, the drivable region and the non-drivable region. The drivable region includes roads, lane lines, and sidewalks, which correspond to roads in agricultural environments.

The threshold map label of the DB Head needs to be calculated according to the edge of the drivable region. Specifically, the Vatti clipping algorithm [34] is used to offset the road edge polygons inside and outside by D, respectively.

$$D = \frac{A(1 - r^2)}{L}$$

where A is the area of the road region polygon, L is the perimeter of the polygon, and $r = 0.4$ is the shrink ratio. The threshold map gradually decreases from a maximum threshold of 0.7 at the edge to 0.3.

In the end, the dataset has 560 labeled samples and 700 RTK samples, of which 1200 are in the training set and 60 in the validation set, as shown in Figure 3.

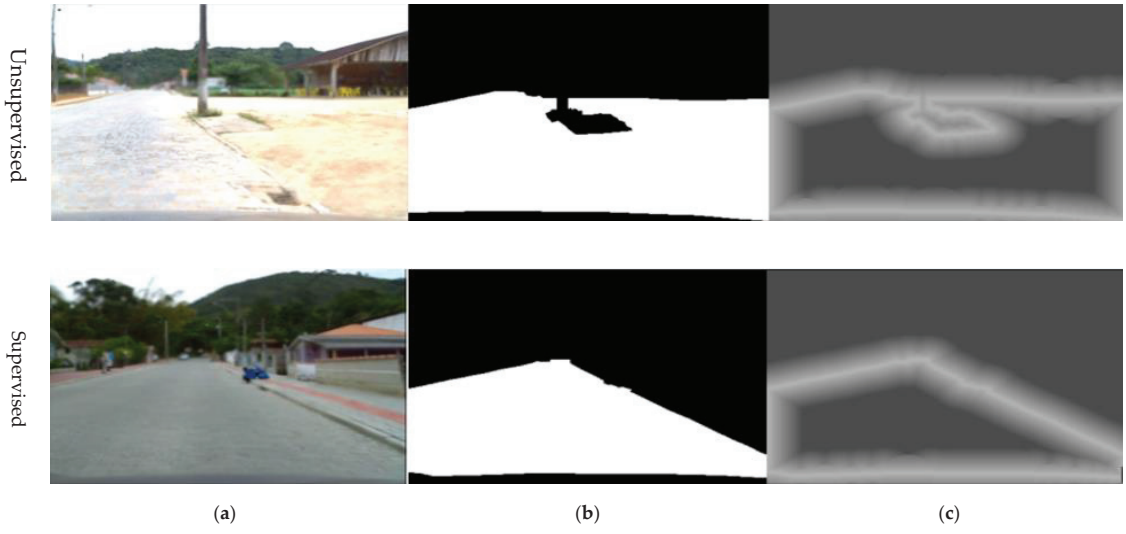


Figure 3. Examples of the agricultural semi-structured road dataset. (**a**) shows the input figure; (**b**) shows the mask for the road region; (**c**) shows the binary mask for the threshold map.

3.2. Evaluation Metrics

In this paper, the mean intersection of union (*mIoU*) of segmentation quality and *Boundary IoU* of boundary quality are selected as the primary evaluation criteria, and pixel accuracy (*PA*) is supplemented as the basis for comparison.

The *mIoU* is a common metric to measure the image overlap rate in object detection. The larger the average cross-union ratio, the more accurate the target prediction is. The calculation formula for *mIoU* is

$$mIoU = \frac{1}{k+1} \sum_{i=0}^{k} \frac{p_{ii}}{\sum_{j=0}^{k} p_{ij} + \sum_{j=0}^{k} p_{ji} - p_{ii}}$$

where k is the number of classifications in the segmentation task, p_{ii} is the i class label predicted to be the i class pixel, p_{ij} is the i class label predicted to be the j class pixel.

Compared with a large number of simple samples, the proportion of boundary samples prone to ambiguity is small. *mIoU* cannot accurately reflect the boundary quality of the segmentation algorithm. We adopt *Boundary IoU* [35] to measure the boundary segmentation accuracy of the segmentation algorithm.

$$Boundary\ IoU = \frac{|a|}{|(G_d \cap G) \cup (P_d \cap P)|}$$

where G is the ground truth binary mask, P is the prediction binary mask, G_d and P_d are the set of pixels in the boundary region of the binary mask.

The PA is the ratio of the number of correctly predicted pixels to the total number of pixels, which is calculated as:

$$PA = \frac{\sum_{i=0}^{k} p_{ii}}{\sum_{i=0}^{k} \sum_{j=0}^{k} p_{ij}}$$

where k is the number of classifications in the segmentation task. This paper includes the drivable region class and non-drivable region class, where the drivable region is class 1, and the non-drivable region class is class 0.

3.3. Implementation and Training Details

We conduct experiments using PyTorch 2.0 on a personal workstation with two NVIDIA GTX 2080Ti. To fully exploit the expressive ability of the model, we use a series of data enhancements, including mirror flip, top-down flip, z-axis flip, and random scaling in the range of [0.5, 2]. We used AdamW as the optimizer for the training strategy, setting the learning rate to 0.0001.

4. Results and Discussion

4.1. Experimental Results and Comparison

4.1.1. Detection Results on the Dataset

We compare our method with DeepLabV3+ [26] and HRNet+OCR [29], which are state-of-the-art methods. DeepLabV3+ is used as the baseline for quantification. The recognition performance of different algorithms for agricultural semi-structured roads is shown in Table 1. The *mIoU* of our method is 97.85%, and the *Boundary IoU* is 90.88%. Compared with DeepLabV3+, the *mIoU* is improved by 2.73%, and the *Boundary IoU* is improved by 5.52%. Compared with HRNet+OCR, the *mIoU* is improved by 0.72%, and the *Boundary IoU* is improved by 2.92%. This result shows that our method has good performance on the dataset.

Table 1. Segmentation results of different methods for agricultural semi-structured roads.

Method	*mIoU* (%)	*Boundary IoU* (%)	Comparison of *mIoU* (%)	Comparison of *Boundary IoU* (%)
DeepLabV3+ (Baseline)	95.12	85.36	0	0
HRNet+OCR	97.13	87.96	2.01%	2.6
Ours	97.85	90.88	2.73%	5.52

4.1.2. Detection Results in Agricultural Scenarios

To further test the practical performance of the method, we conduct real-time semi-structured road region detection experiments on the farm. Test sites include indoor and outdoor sites. The indoor test site is a corridor inside a greenhouse. Lush crops are planted around the corridors that obscure the borders of the road. Meanwhile, crops block the sun, leaving dappled shadows on the road. The outdoor experimental site is a cement road in the lawn outside the greenhouse. Lawn grows in clutter, with no clear boundaries between roads and lawns. Figure 4 visually presents the segmentation effect of different methods on agricultural semi-structured roads. Although DeepLabV3+, HRNet+OCR and our method can detect drivable regions, the detection accuracy is different. The detection performance of our method outperforms the other two methods both indoors and outdoors. Especially in the presence of disturbances, our method is more robust. Specifically, in scene 1, a crack on the road surface affects the detection accuracy of DeeplabV3+ and HRNet+OCR. In scene 2, DeeplabV3+ and HRNet+OCR cannot segment distant roads because the grass and road features are similar. The large amount of shadows in scene 3 destroys the visual characteristics of the road, making the traditional segmentation method invalid. In contrast,

our method can achieve robust detection of semi-structured roads in complex scenes. The field test results are consistent with the results on the dataset.

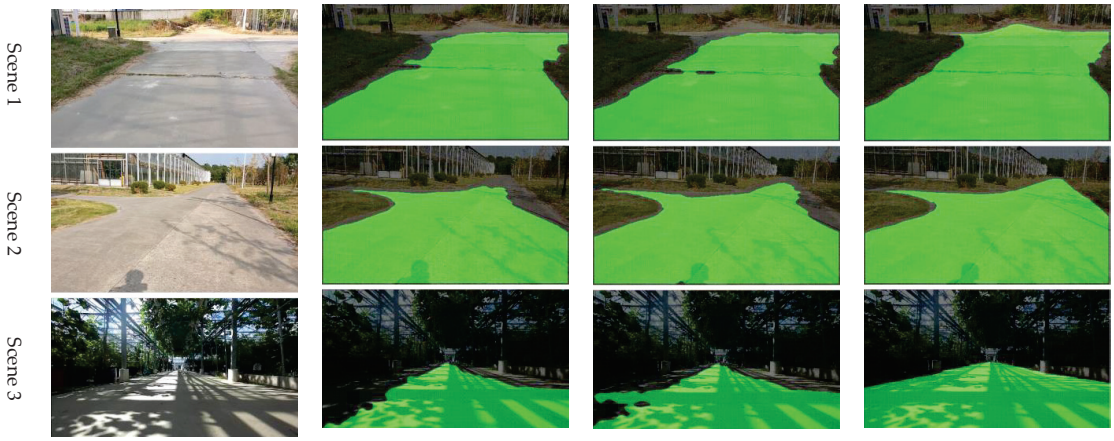


Figure 4. Examples of segmentation results for agricultural semi-structured roads. (**a**) shows the input image; (**b**) shows the result by the DeepLab V3+ (baseline); (**c**) shows the result by HRNet+OCR; (**d**) shows the result by this thesis.

4.2. Discussion

4.2.1. The Effect of Threshold Map Supervision on Road Detection

We first discuss the effect of adding threshold map supervision in the DB module on road detection performance. We compare road detection accuracy with and without threshold map supervision on the dataset. The experimental results are shown in Table 2. The results demonstrate that using threshold map supervision, the *mIoU* of the method is improved by 0.48%, the *PA* is improved by 0.17%, and the *Boundary IoU* is improved by 2.17%. Supervising the threshold map can significantly improve the quality of boundary detection.

Table 2. Comparison results of whether the threshold map is supervised.

Method	*mIou* (%)	*PA* (%)	*Boundary IoU* (%)
Unsupervised	97.16	98.64	88.71
Supervised	97.64	98.81	90.88
Comparison of results	0.48	0.17	2.17

Figure 5 shows the boundary quality improvement by the threshold map supervision. Compared with the unsupervised case, the threshold map with supervision can learn more accurate road boundaries. Especially when there are disturbances at the boundaries, the threshold map without supervision easily identifies disturbances as the road boundaries. For example, Figure 5b misidentifies water spots and steps as road boundaries, which damages the segmentation accuracy of the road in Figure 5c. In contrast, the threshold map with supervision can learn correct boundary thresholds on similar image features, resulting in sharp boundaries.

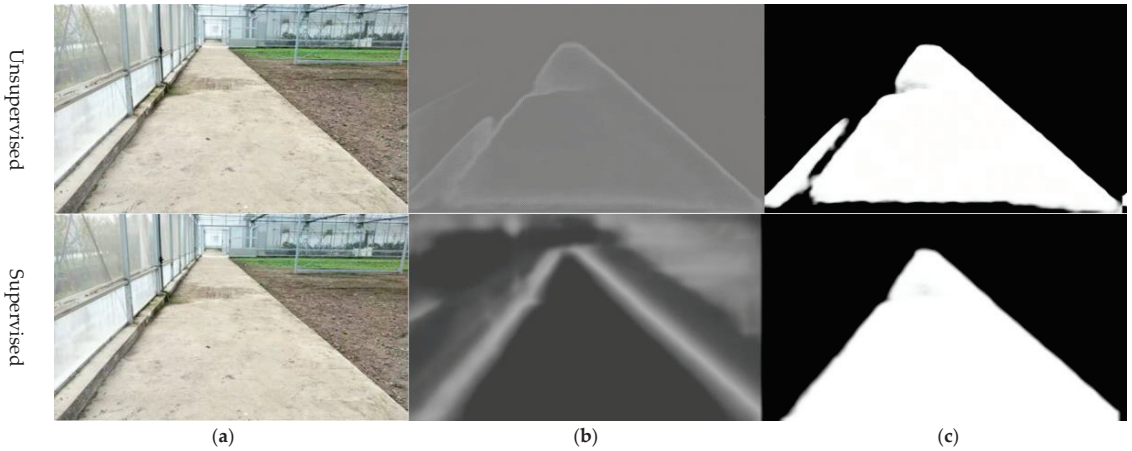


Figure 5. Comparison of segmentation results with threshold map supervision. (**a**) shows the input image; (**b**) shows the result by the boundary threshold; (**c**) shows the result by the segmentation result.

4.2.2. Comparative Analysis of Different Loss Functions

Based on threshold supervision, we compare loss functions that balance between positive and negative samples as well as easy and hard samples. In supervising coarse features, this study utilizes the BCE loss function rather than the hard negative mining method. This ensures that coarse features retain a normal distribution, allowing the entire OCR module to play a role in feature fine-tuning. For the probability map and the approximate binary map, we employ the hard negative mining approach.

Table 3 shows the impact of different hard negative mining methods on the network accuracy. The data show that the performance of WCE is superior to that of BCE. This is attributed to the fact that WCE balances the loss between positive and negative samples by assigning weights based on the dataset's statistical distribution of these samples. Meanwhile, focal loss achieves better results due to its consideration of balancing between easy and hard samples. Pixels in the boundary regions are given higher weights during the training process. In addition, OHEM delivers the best performance. It effectively balances between easy and hard samples by sorting and filtering the scores of the negative samples. Compared to the baseline model, there was an improvement of 0.63% in *Road IoU* and a 1.56% increase in *Boundary IoU*. This indicates its efficacy on facility agricultural roads with ambiguous boundaries, enhancing the quality of road boundary segmentation. This is crucial for subsequent tasks such as edge localization [36,37].

Table 3. The effect of different difficult sample mining methods on accuracy.

Loss Function	*Road IoU* (%)	*Boundary IoU* (%)
BCE (Baseline)	97.22	89.32
WCE	97.38	89.60
Focal loss	97.50	89.92
OHEM	97.85	90.88

4.2.3. Discussion of Different Deployment Methods

We tested the performance of different deployment methods on the dataset. These deployment methods are single-precision floating point (FP32), half-precision floating point (FP16), and data enhancements (including mirror flip, top-down flip, z-axis flip, and random scaling). Table 4 shows the results obtained on the dataset with different deployment methods.

Table 4. Experimental results of different deployment methods.

Deployment Method	*mIoU* (%)	*PA* (%)	*Boundary IoU* (%)
FP32 (Baseline)	98.73	98.81	90.88
FP16	98.72	98.80	90.74
Data enhancements	97.78	98.87	91.30

When the network parameters use the half-precision floating point, the *mIoU* is reduced by 0.01% compared with using the single-precision floating point, and the *Boundary IoU* is only reduced by 0.14%. This means that using the half-precision floating point for network parameters does not affect road detection accuracy, while reducing inference time by nearly 33%. Therefore, selecting the half-precision floating point for network parameters is recommended during training. Using data enhancements can slightly improve the accuracy of road detection, but it will increase the training time.

5. Conclusions

This paper proposes a high-resolution detection network for semi-structured road segmentation for autonomous navigation of agricultural robots. Since the ambiguous road boundaries cannot be semantically identified with the contrast against their neighborhoods, the pixel-level contextual representations are used at different scales to augment the road objectness. To extract high-resolution features of semi-structured roads, the HRNet is adopted as the backbone to preserve high-resolution features during downsampling. To distinguish the features of drivable and non-drivable regions more clearly, we can use a relational context-based OCR module to enhance the feature representations. Finally, an adaptive threshold-learned DB decision head is used to segment road boundaries.

To quantify the performance of our method, we used an agricultural semi-structured road dataset and conducted experiments. The experimental results show that the *mIoU* reaches 97.85%, and the *Boundary IoU* achieves 90.88%. The segmentation accuracy and boundary quality are better than the existing methods, which can meet the needs for the accurate detection of semi-structured drivable areas.

Author Contributions: Conceptualization, Y.S., L.G. and W.Z.; data curation, L.G. and W.Z.; formal analysis, C.L.; funding acquisition, L.G.; investigation, L.G.; methodology, Y.S., B.G. and Y.L.; project administration, L.G.; software, Y.S. and W.Z.; supervision, C.L.; validation, B.G.; visualization, Y.S. and B.G.; writing—original draft, Y.S., L.G. and W.Z.; writing—review and editing, Y.S. and L.G. All authors have read and agreed to the published version of the manuscript.

Funding: This work was supported by Shanghai Agriculture Applied Technology Development Program, China (Grant No. T2022-03-01) and Shanghai Science and Technology Committee (STCSM) Science and Technology Innovation Program (Grant No. 20392000900).

Institutional Review Board Statement: The study did not require ethical approval.

Data Availability Statement: The data cannot be made publicly available upon publication because they are owned by a third party and the terms of use prevent public distribution. The data that support the findings of this study are available upon reasonable request from the authors.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Zhang, W.; Gong, L.; Sun, Y.; Gao, B.; Yu, C.; Liu, C. Precise visual positioning of agricultural mobile robots with a fiducial marker reprojection approach. *Meas. Sci. Technol.* **2023**, *34*, 115110. [CrossRef]
2. Bechar, A.; Vigneault, C. Agricultural robots for field operations: Concepts and components. *Biosyst. Eng.* **2016**, *149*, 94–111. [CrossRef]
3. Qi, N.; Yang, X.; Chuanxiang, L.; Lu, R.; He, C.; Cao, L. Unstructured Road Detection via Combining the Model-based and Feature-based Methods. *IET Intell. Transp. Syst.* **2019**, *13*, 1533–1544. [CrossRef]
4. Xiao, L.; Dai, B.; Liu, D.; Zhao, D.; Wu, T. Monocular Road Detection Using Structured Random Forest. *Int. J. Adv. Robot. Syst.* **2016**, *13*, 101. [CrossRef]

5. Yang, G.; Wang, Y.; Yang, J.; Lu, Z. Fast and Robust Vanishing Point Detection Using Contourlet Texture Detector for Unstructured Road. *IEEE Access* **2019**, *7*, 139358–139367. [CrossRef]
6. Shi, J.; Wang, J.; Fu, F. Fast and Robust Vanishing Point Detection for Unstructured Road Following. *IEEE Trans. Intell. Transp. Syst.* **2016**, *17*, 970–979. [CrossRef]
7. Liu, Y.B.; Zeng, M.; Meng, Q.H. Unstructured Road Vanishing Point Detection Using Convolutional Neural Networks and Heatmap Regression. *IEEE Trans. Instrum. Meas.* **2021**, *70*, 1–8. [CrossRef]
8. Hernandez, D.E.; Blumenthal, S.; Prassler, E.; Bo, S.; Haojie, Z. Vision-based road boundary tracking system for unstructured roads. In Proceedings of the 2017 IEEE International Conference on Unmanned Systems (ICUS), Beijing, China, 27–29 October 2017; pp. 66–71.
9. Liu, G.; Worgotter, F.; Markelic, I. Stochastic Lane Shape Estimation Using Local Image Descriptors. *IEEE Trans. Intell. Transp. Syst.* **2013**, *14*, 13–21. [CrossRef]
10. Perng, J.W.; Hsu, Y.W.; Yang, Y.Z.; Chen, C.Y.; Yin, T.K. Development of an embedded road boundary detection system based on deep learning. *Image Vis. Comput.* **2020**, *100*, 103935. [CrossRef]
11. Cao, J.; Song, C.; Song, S.; Xiao, F.; Peng, S. Lane Detection Algorithm for Intelligent Vehicles in Complex Road Conditions and Dynamic Environments. *Sensors* **2019**, *19*, 3166. [CrossRef]
12. Wang, K.; Huang, Z.; Zhong, Z. Algorithm for urban road detection based on uncertain Bezier deformable template. *Jixie Gongcheng Xuebao Chin. J. Mech. Eng.* **2013**, *49*, 143–150. [CrossRef]
13. Yuan, Y.; Jiang, Z.; Wang, Q. Video-based road detection via online structural learning. *Neurocomputing* **2015**, *168*, 336–347. [CrossRef]
14. Xiang, W.; Juan, Z.; Zhijun, F. Unstructured road detection based on contour selection. In Proceedings of the 4th International Conference on Smart and Sustainable City (ICSSC 2017), Shanghai, China, 5–6 June 2017; pp. 1–5.
15. Li, J.; Liu, C. Research on Unstructured Road Boundary Detection. In Proceedings of the 2021 IEEE International Conference on Unmanned Systems (ICUS), Beijing, China, 15–17 October 2021; pp. 614–617.
16. Alam, A.; Singh, L.; Jaffery, Z.A.; Verma, Y.K.; Diwakar, M. Distance-based confidence generation and aggregation of classifier for unstructured road detection. *J. King Saud Univ.—Comput. Inf. Sci.* **2022**, *34*, 8727–8738. [CrossRef]
17. Sturgess, P.; Alahari, K.; Ladicky, L.; Torr, P. Combining Appearance and Structure from Motion Features for Road Scene Understanding. In Proceedings of the British Machine Vision Conference, BMVC 2009, London, UK, 7–10 September 2009.
18. Baheti, B.; Innani, S.; Gajre, S.; Talbar, S. Eff-UNet: A Novel Architecture for Semantic Segmentation in Unstructured Environment. In Proceedings of the 2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops (CVPRW), Seattle, WA, USA, 14–19 June 2020; pp. 1473–1481.
19. Wang, Q.; Fang, J.; Yuan, Y. Adaptive road detection via context-aware label transfer. *Neurocomputing* **2015**, *158*, 174–183. [CrossRef]
20. Geng, L.; Sun, J.; Xiao, Z.; Zhang, F.; Wu, J. Combining CNN and MRF for road detection. *Comput. Electr. Eng.* **2018**, *70*, 895–903. [CrossRef]
21. Ronneberger, O.; Fischer, P.; Brox, T. U-net: Convolutional networks for biomedical image segmentation. In Proceedings of the Medical Image Computing and Computer-Assisted Intervention–MICCAI 2015: 18th International Conference, Munich, Germany, 5–9 October 2015; Proceedings, Part III 18, 2015. pp. 234–241.
22. Badrinarayanan, V.; Kendall, A.; Cipolla, R. SegNet: A Deep Convolutional Encoder-Decoder Architecture for Image Segmentation. *IEEE Trans. Pattern Anal. Mach. Intell.* **2017**, *39*, 2481–2495. [CrossRef]
23. Chen, L.C.; Papandreou, G.; Kokkinos, I.; Murphy, K.; Yuille, A.L. DeepLab: Semantic Image Segmentation with Deep Convolutional Nets, Atrous Convolution, and Fully Connected CRFs. *IEEE Trans. Pattern Anal. Mach. Intell.* **2018**, *40*, 834–848. [CrossRef]
24. Chen, L.-C.; Papandreou, G.; Kokkinos, I.; Murphy, K.; Yuille, A.L. Semantic image segmentation with deep convolutional nets and fully connected crfs. *arXiv* **2014**, arXiv:1412.7062.
25. Chen, L.-C.; Papandreou, G.; Schroff, F.; Adam, H. Rethinking atrous convolution for semantic image segmentation. *arXiv* **2017**, arXiv:1706.05587.
26. Chen, L.-C.; Zhu, Y.; Papandreou, G.; Schroff, F.; Adam, H. Encoder-Decoder with Atrous Separable Convolution for Semantic Image Segmentation. In Proceedings of the European Conference on Computer Vision (ECCV), Munich, Germany, 8–14 September 2018; pp. 801–818, 833–851.
27. Li, J.; Jiang, F.; Yang, J.; Kong, B.; Gogate, M.; Dashtipour, K.; Hussain, A. Lane-DeepLab: Lane semantic segmentation in automatic driving scenarios for high-definition maps. *Neurocomputing* **2021**, *465*, 15–25. [CrossRef]
28. Wang, J.; Sun, K.; Cheng, T.; Jiang, B.; Deng, C.; Zhao, Y.; Liu, D.; Mu, Y.; Tan, M.; Wang, X.; et al. Deep High-Resolution Representation Learning for Visual Recognition. *IEEE Trans. Pattern Anal. Mach. Intell.* **2021**, *43*, 3349–3364. [CrossRef]
29. Yuan, Y.; Chen, X.; Wang, J. Object-contextual representations for semantic segmentation. In Proceedings of the Computer Vision–ECCV 2020: 16th European Conference, Glasgow, UK, 23–28 August 2020; Proceedings, Part VI 16, 2020. pp. 173–190.
30. Liao, M.; Zou, Z.; Wan, Z.; Yao, C.; Bai, X. Real-Time Scene Text Detection With Differentiable Binarization and Adaptive Scale Fusion. *IEEE Trans. Pattern Anal. Mach. Intell.* **2023**, *45*, 919–931. [CrossRef]
31. Vaswani, A.; Shazeer, N.; Parmar, N.; Uszkoreit, J.; Jones, L.; Gomez, A.N.; Kaiser, Ł.; Polosukhin, I. Attention is all you need. *Adv. Neural Inf. Process. Syst.* **2017**, *30*. [CrossRef]

32. Hao, Y.; Liu, Y.; Wu, Z.; Han, L.; Chen, Y.; Chen, G.; Chu, L.; Tang, S.; Yu, Z.; Chen, Z. Edgeflow: Achieving practical interactive segmentation with edge-guided flow. In Proceedings of the IEEE/CVF International Conference on Computer Vision, Virtual, 11–17 October 2021; pp. 1551–1560.
33. Rateke, T.; Justen, K.A.; Von Wangenheim, A. Road surface classification with images captured from low-cost camera-road traversing knowledge (rtk) dataset. *Rev. Inf. Teórica Apl.* **2019**, *26*, 50–64. [CrossRef]
34. Vatti, B.R. A generic solution to polygon clipping. *Commun. ACM* **1992**, *35*, 56–63. [CrossRef]
35. Cheng, B.; Girshick, R.; Dollár, P.; Berg, A.C.; Kirillov, A. Boundary IoU: Improving object-centric image segmentation evaluation. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, Nashville, TN, USA, 19–25 June 2021; pp. 15334–15342.
36. Sun, Y. A Comparative Study on the Monte Carlo Localization and the Odometry Localization. In Proceedings of the 2022 IEEE International Conference on Electrical Engineering, Big Data and Algorithms (EEBDA), Changchun, China, 25–27 February 2022; pp. 1074–1077.
37. Zhang, W.; Gong, L.; Huang, S.; Wu, S.; Liu, C. Factor graph-based high-precision visual positioning for agricultural robots with fiducial markers. *Comput. Electron. Agric.* **2022**, *201*, 107295. [CrossRef]

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.

agriculture

MDPI

Article

Optimal Sand—Paving Parameters Determination of an Innovatively Developed Automatic Maize Seeding Machine

Bohan Fu [1,2,3], Weizhong Sun [1,2,3] and Zhao Zhang [1,2,3,4,*]

[1] Key Laboratory of Smart Agriculture System Integration, Ministry of Education, Beijing 100083, China; 2020308130329@cau.edu.cn (B.F.); 2020308130208@cau.edu.cn (W.S.)
[2] Key Laboratory of Agricultural Information Acquisition Technology, Ministry of Agriculture and Rural Affairs, China Agricultural University, Beijing 100083, China
[3] College of Information and Electrical Engineering, China Agricultural University, Beijing 100083, China
[4] Department of Agricultural and Biosystems Engineering, North Dakota State University, Fargo, ND 58102, USA
* Correspondence: zhaozhangcau@cau.edu.cn

Abstract: Maize is an important crop to ensure food safety. High-quality seeds can guarantee a good yield. The maize seed germination rate is the most important information for the maize industry, which can be obtained through the seed germination test. An essential stage in determining the germination rate is the planting of the seeds. The current seed planting process is fully manual, which is labor-intensive and costly, and it requires the development of an autonomous seeding machine. This research developed an automatic maize seeding machine, consisting of four operations: paving sand, seed layout, watering, and covering the seed. Among the four procedures, sand paving is a crucial step, the performance of which is affected by the gate opening size, conveyor speed, and sensor mounting location. Three performance evaluating factors are the weight of sand in the tray, the volume of sand left on the conveyor, and sand surface flatness. A full factorial experiment was designed with three variables and three levels to determine an appropriate factor combination. RGB-D information was used to calculate the volume of sand left on the conveyor and sand flatness. An analytic hierarchy process was employed to assign weights to the three evaluation indicators and score the various combinations of factors. The machine for paving sand achieved a satisfactory result with an opening size of 10.8 mm, a sensor distance of 9 cm, and a conveyor belt speed of 5.1 cm/s. With the most satisfactory factors determined, the machine shows superior performance to better meet practical applications.

Keywords: automatic maize seeding machine; sand-paving device; image processing; analytic hierarchy process

Citation: Fu, B.; Sun, W.; Zhang, Z. Optimal Sand—Paving Parameters Determination of an Innovatively Developed Automatic Maize Seeding Machine. *Agriculture* **2023**, *13*, 1538. https://doi.org/10.3390/agriculture13081538

Academic Editors: Jin Yuan and Zichen Huang

Received: 9 July 2023
Revised: 26 July 2023
Accepted: 28 July 2023
Published: 2 August 2023



Copyright: © 2023 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (https://creativecommons.org/licenses/by/4.0/).

1. Introduction

Maize production plays a crucial role in ensuring global food security [1], and China holds the position of the second-largest maize producer worldwide [2]. The germination rate of seeds holds significant importance as it directly impacts the overall yield, serving as a vital indicator of seed quality. In China's seed production industry, Standard GB/T 5243.4 is employed to determine the seed germination rate, encompassing both seed planting and seedling counting processes. While computer vision has been utilized in various studies for automated seedling counting [3,4], seed planting remains a fully manual process. The manual sowing procedure involves several distinct steps. The process starts with the operator preparing the tray. Firstly, they cover the tray with sand and then flatten it to create a smooth surface (Step 1). Next, the operator evenly distributes the seeds onto the sandy surface of the tray. To aid germination, they spray water onto the seeds (Step 2). In order to ensure optimal conditions for seed growth, another layer of sand is then added on top of the seeds, providing additional protection and support (Step 3). Finally, the

prepared tray is placed into an incubator that creates a suitable environment to promote seed germination (Step 4). The subjective nature of operators can have an impact on the results of the experiment [5–7]. Moreover, the repetitive and extended tasks of sand flattening and sowing can pose occupational health risks for workers, including back and wrist injuries [8]. Therefore, there is a need for the development of an autonomous maize seeding machine.

This study designed and developed a maize seed sowing machine to automatize the seeding process, which encompasses four modules: sand-paving tray, seed release mechanism, water spraying, and covering the seed with sand. The initial step involves paving the tray with sand, followed by the crucial task of leveling or flattening the sand. Without proper flattening, when seeds are released onto the sand, there is a risk of them rolling away from their intended position, disrupting the predetermined uniform pattern. This disruption can lead to challenges in accurately counting and assessing the stand, as multiple seeds may germinate in close proximity. Therefore, maintaining a flat sand surface is essential, as it ensures the desired uniformity for successful seed distribution and germination.

The levelness of the sand surface directly influences seed germination. In a study by researchers [9], different soil-covering and compacting devices were compared, and factors such as soil morphology and compaction were evaluated. Xu et al. (2021) developed a mathematical model to describe the relationship between soil output, mulch disc depth, and mulch disc helical blades [10]. To overcome the subjectivity and inaccuracy associated with manual observation, this experiment employs an RGB-D camera to quantitatively assess sand volume and flatness [11]. It has been demonstrated by Zhou et al. (2021) that RGB-D cameras can effectively capture RGB and depth data [12]. These cameras have gained wide adoption in agriculture and have shown favorable results in various practical applications, including plant phenotyping [13], 3D segmentation of plants, and obstacle detection for agricultural robots [14,15]. Therefore, utilizing an RGB-D camera for evaluating the sand surface represents a promising option.

With the ultimate objective of achieving practical implementation of the automatic maize seeding machine, this study aimed to explore the impact of three factors, namely gate opening size, sensor mounting location, and conveyor speed, on the evaluation of sand surface flatness. The goal was to determine the optimal combination of parameters for effective practical application.

2. Materials and Methods

2.1. Brief Introduction to the Automatic Maize Seeding Machine

The automatic maize seeding machine, shown in Figure 1, has been innovatively designed. It comprises a conveyor belt and four distinct modules operating independently: (1) sand paving, (2) seed placement, (3) watering, and (4) sand covering. At the beginning of the process, an empty tray is positioned on one end of the machine (the left-hand side in Figure 1). As the conveyor belt transports the tray forward, its position is constantly monitored by sensors (GP2Y0A21, manufactured by Sharp Corporation in Osaka, Japan). When the tray is detected nearing the sand-paving device, the motor is activated, opening the gate to commence the sand-paving process. Upon detection of the tray leaving the sand-paving device, the motor is triggered to close the gate, signaling the end of the sand-paving process. Following the sand-paving phase, the tray proceeds to the seed placement mechanism, where seeds are released onto the sand surface. Subsequently, the tray moves forward into the watering section. Finally, the tray reaches the second sand-filling section, where the seeds are covered with sand.

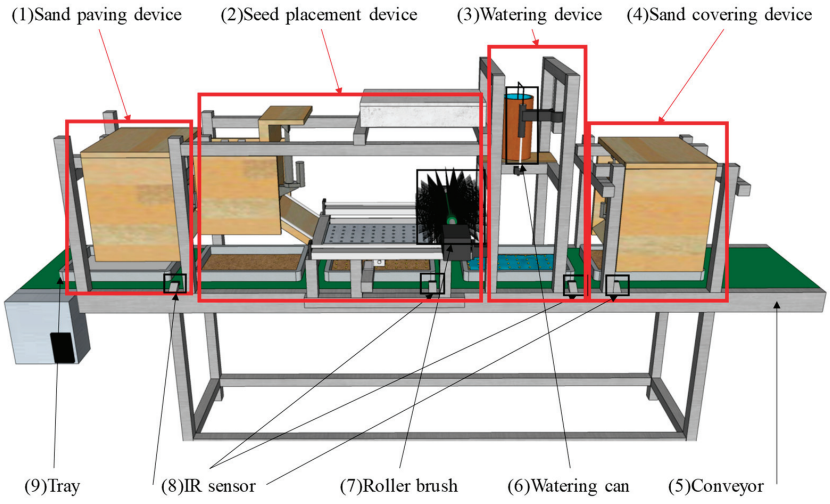


Figure 1. Automatic maize seeding apparatus: (1) Sand-paving device; (2) Seed placement device; (3) Watering device; (4) Sand-covering device; (5) Conveyor for transporting trays; (6) Watering can; (7) Roller brush for sweeping the seeds; (8) IR sensor for monitoring tray location; and (9) Empty tray for seed germination test.

The automatic maize seeding machine was developed by our team to fully replace the labor for seeding. Among the four key procedures (i.e., sand paving, seed placement, watering, and sand covering), the first procedure of sand paving is a crucial step, which functions to pave the tray with sand. The evenness of the sand surface is crucial for successful sand paving. Uneven surfaces can cause seed displacement and result in seeds rolling away from their intended landing location. This can lead to inconsistencies in the final location of the seeds and negatively impact the overall quality of the paving. The sand-paving device plays a crucial role in determining the machine's performance. Prior to the seeds being released onto the sand, they are arranged in a desirable uniform pattern. Any disruption to this pattern can result in two seeds germinating at the same location, presenting a challenge when it comes to accurately counting the seedlings. Therefore, the efficacy of the sand-paving device is critical for ensuring successful seed distribution and subsequent seedling monitoring.

2.2. Sand-Paving Device

The sand-paving device (Figure 2) has a dimension of 297 mm × 290 mm × 400 mm, and the bottom plate has a slope of 45°, which would automatically move the sand downward to the tray. The gate open and close is controlled by a 24V DC motor with a max 180 mm/s linear speed and a 100 mm stroke. The control system of the paving device utilizes an Arduino board (UNO R3, Arduino, Strambino, Italy). The system operates as follows: the sensor detects the presence of the tray by measuring the distance between the sensor and the tray. When the tray enters the sensor's view, the measured distance decreases from over 10 cm to less than 10 cm. This prompts the Arduino to send a signal to the relays (JQC-3FF-S-Z 5V D.C.; Shenzhen Weixin, Shenzhen, China) in order to open the gate. Conversely, when the tray moves out of the sensor's view, the measured distance increases from less than 10 cm to over 10 cm, signaling the Arduino to close the gate by triggering the relays.

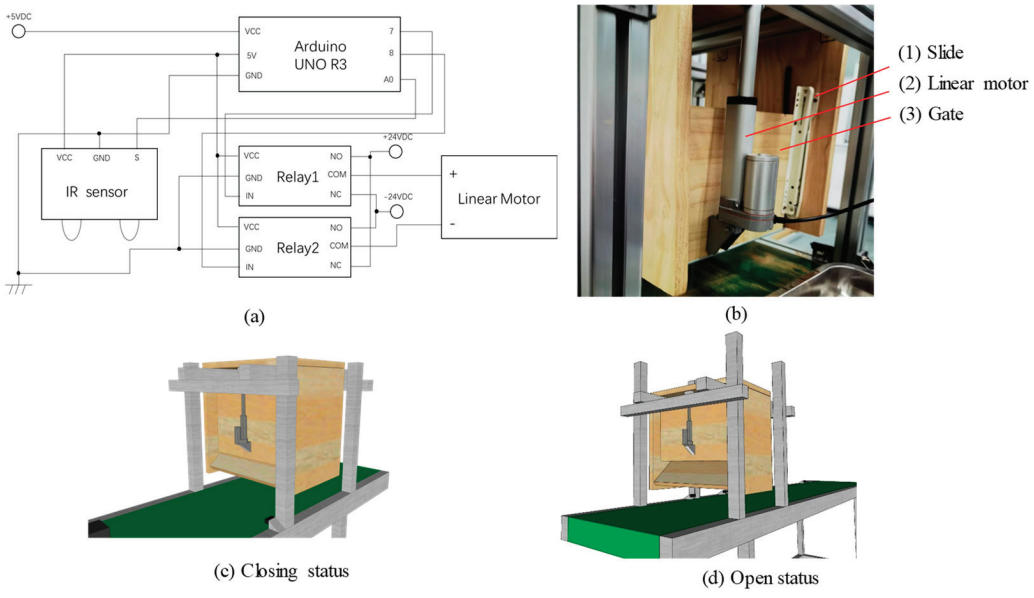


Figure 2. Newly developed sand−paving device: (**a**) Circuit diagram; (**b**) Sand-paving device; (**c**) gate in closed mode; and (**d**) gate in open mode.

2.3. Experimental Procedure

A full factorial experiment design was employed to evaluate the significance of factors and explore all possible combinations of factor levels. This approach enables a comprehensive comparison of the impact of different factor combinations on machine performance. The first factor examined is the conveyor speed, which influences the duration of the tray's movement beneath the sand gate and consequently affects the flatness of the sand surface. The second factor is the gate opening size, which directly determines the rate at which sand is filled. Lastly, the relative distance of the sensor to the sand-filling gate is examined as the third factor.

The three factors were configured with three levels each: (1) conveyor speeds of 0.13, 0.51, and 0.65 m/s; (2) gate opening sizes of 7.2, 9.0, and 10.8 mm; and (3) sensor positions relative to the sand control gate of 40.0, 70.0, and 90.0 mm (the relative distances were determined through preliminary experiments). In total, there were 27 unique settings, with each setting replicated three times, resulting in a total of 81 runs.

After each run, three parameters were recorded: the weight of the tray, the volume of sand remaining on the conveyor belt, and the flatness of the sand surface. Additionally, RGB-D information was collected using an Intel RealSense D435i camera from Santa Clara, California, U.S. These data were collected for both the tray and the sand on the conveyor belt. To obtain the tray weight, the tray was carefully moved away from the conveyor and weighed using a scale (SNJ-50001; SINUOJIE, Shenzhen, China). The image data were processed using MATLAB R2022a (The Mathworks, Inc., Natick, MA, USA). Origin 2022 (OriginLab, Inc., Northampton, Mass., USA) was used for image plotting, while third-party libraries such as NumPy and OpenCV were used in Python. Figure 3 shows the overall procedure for determining the satisfactory combination of factors.

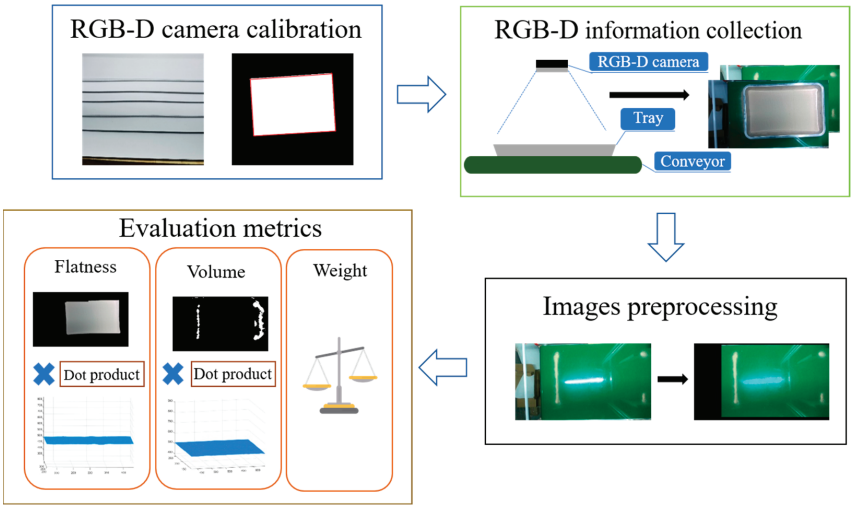


Figure 3. The overall procedure of determining the optimal sand-paving parameters of a newly developed automatic maize seeding machine.

2.4. Calibration of the RGB-D Data

2.4.1. Depth Information

The evenness of the sand surface is one of the more important topics in this research. However, directly measuring the sand surface presents challenges. In this study, RGB-D information is leveraged to measure the flatness. Before utilizing the depth information to evaluate sand evenness, a calibration process is conducted to ensure accurate sensor data.

A manually built staircase using an A4 sheet with known dimensions was used to calibrate the depth data (Figure 4). In this study, the effect of ambient light is not received by the calibration effect, as these data were collected in a dark environment. Twenty A4 sheets were used to construct each step of the stairs, and their thickness was maintained at 3 mm after measurement with a caliper. It was found that the measuring error was within 0.1 mm. Figure 4 visually demonstrates the evident staircase pattern, with a noticeable 3 mm difference in height between steps. This suggests that the depth data resolution can achieve a level of 1 mm, making it suitable for assessing the evenness of the sand surface.

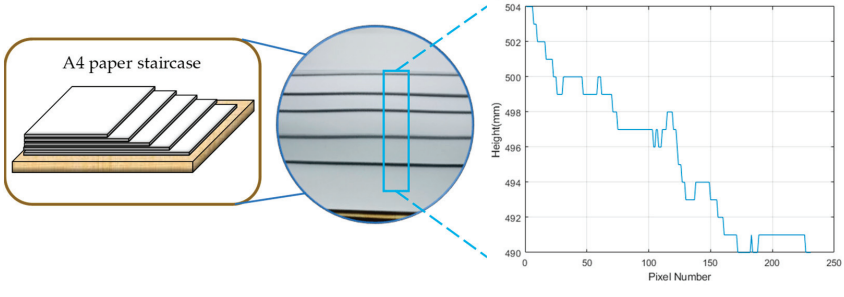


Figure 4. RGB-D camera calibration for depth information using A4 paper staircase.

2.4.2. RGB Image Calibration

In addition to calibrating depth information, it is necessary to check the distortion of the collected RGB images [16]. An A4 sheet was placed on the conveyor belt, and an RGB image was collected (Figure 5). A MATLAB® tool, color thresholder, was used to separate the A4 paper and extract the mask image.

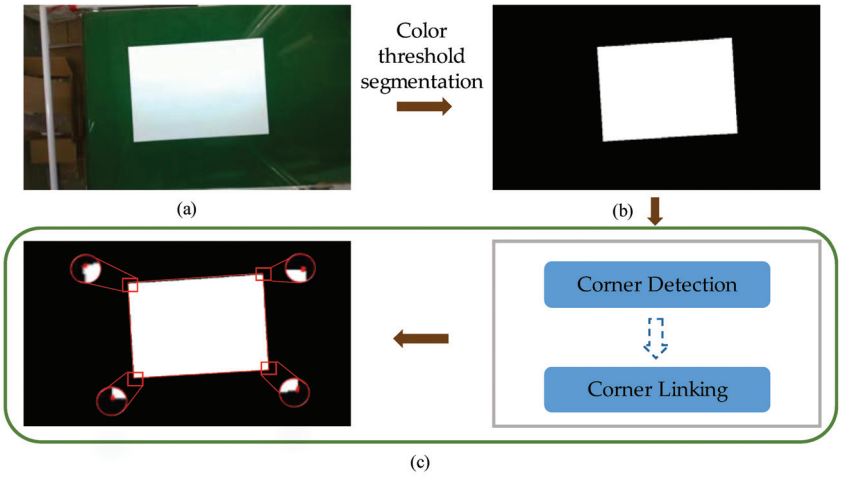


Figure 5. RGB image and segmentation: (**a**) Original image; (**b**) corresponding binary image after thresholding; and (**c**) four vertices of the binary image connected.

The corner Harris function, implemented in the OpenCV library, was utilized to detect the coordinates of the four vertices in the binary image. Upon computation, it was determined that the length of the upper and lower edges of the A4 paper matched that of the left and right edges. Additionally, in Figure 5c, the lines connecting the four vertices closely aligned with the boundaries of the A4 paper in the binary image. From these observations, it can be inferred that the camera-captured images exhibited minimal distortion. The actual area corresponding to a pixel can be calculated by counting the number of white pixels in image N (Figure 5c). The area corresponding to pixels was utilized in the following study as the actual volume of sand needed to be measured. The size of an A4 paper is known to be 210 mm × 297 mm. The value of 0.19 mm^2, obtained after calculation, represents the area corresponding to one pixel as determined by Equation (1).

$$s = \frac{N}{210 \times 297} \tag{1}$$

2.5. Data Acquisition for the Experiment

The RGB-D camera was installed 0.51 m above the conveyor, and the illumination was provided using two LED lights (RS-40; Shenzhen Ruisekeji, Shenzhen, China). Before each experimental run, the weight m_1 of the empty tray was measured, and after the sand was filled in the tray, it was transported by the conveyor belt to the view of the RGB-D camera for data acquisition. The tray was then weighed again with the sand to obtain the weight m_2. Thus, the weight of sand is $m_2 - m_1$. All captured RGB images and depth information tables are standardized with a size of 480 × 848 pixels, signifying that each RGB image pixel corresponds to a specific depth value. Lastly, for each trial, two RGB photos were taken: one of the trays and one of the conveyor belts (capturing the sand left on the tray), along with their corresponding depth information.

2.6. Images Preprocessing

In the collected RGB image (Figure 6a), there is significant noise present in the non-conveyor areas. To facilitate subsequent data processing, the noisy regions of the image were converted into black areas. In the original image, reflections caused by the conveyor belt manifest as bright areas in the picture. These white highlights pose a major challenge for target segmentation, as it becomes difficult to distinguish between the sand and the reflection in the S channel of the HSV (hue saturation value) color space (Figure 6b). To

mitigate this interference, an approach was implemented to identify the highlighted areas and replace them with the color values of the adjacent pixels, effectively eliminating the highlights in the target region of the image. Figure 6d illustrates the noticeable difference after the elimination of highlights, where a clear distinction can be observed between reflection areas and the sand in the S channel. The fast marching method image correction approach, as developed by Telea (2004) [17], was employed for this purpose. As seen in Figure 6c, this repair method demonstrates the potential for producing superior results compared to previous repair algorithms.

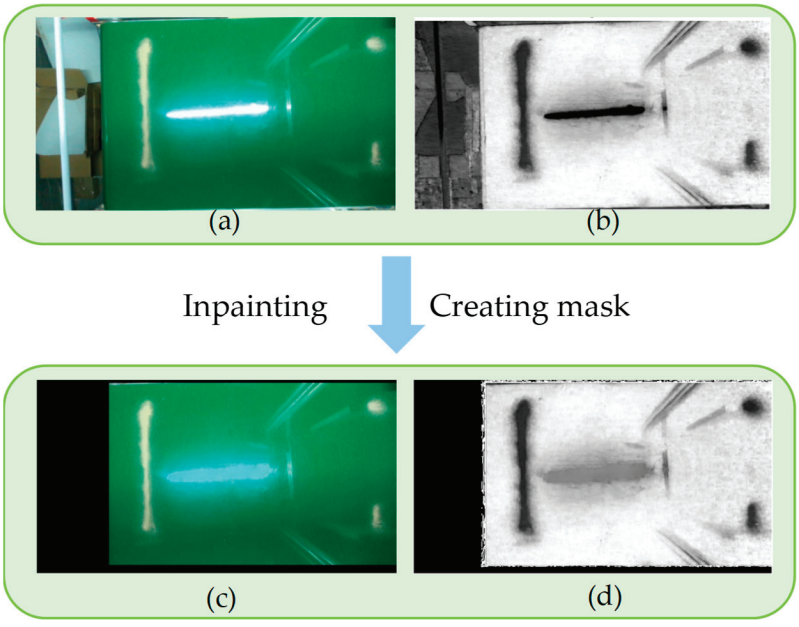


Figure 6. Image preprocessing: (**a**) Original image; (**b**) Original image in HSV color space; (**c**) Image after highlight repair; and (**d**) Image in HSV color space after highlight repair.

2.7. Evaluation Metrics

In this study, the weight of the sand in the tray, the volume of sand on the conveyor (outside the tray), and the evenness of the sand surface are the three parameters employed to determine the optimal combination of factors.

2.7.1. The Weight of Sand in the Tray

The weight of sand plays a crucial role in the sand-filling process and has a significant impact on seed germination [18]. It is directly measured in the experiment using an electronic scale. The weight of sand applied in each experiment is determined by weighing the empty tray and subsequently re-weighing the tray after the sand-paving process is completed. The difference between the two measurements represents the weight of sand in the tray. A higher sand weight generally ensures a favorable environment for seed germination and is therefore preferred.

2.7.2. The Volume of Sand Leaked onto the Conveyor Belt

The amount of sand that spills onto the conveyor belt can vary depending on the timing of the gate opening and the speed at which the conveyor moves during the sand-paving process. Any sand that falls outside the tray is considered a loss and cannot be reused. An increase in the volume of sand outside the tray leads to a proportional increase in the magnitude of loss. Therefore, when dealing with different levels of the selected

factors, measures should be taken to minimize these losses, particularly by reducing the volume of sand that falls outside the tray. The sand on the conveyor belt needs to be separated from the image, and then the heights of each pixel corresponding to the sand should be summed to calculate the amount of sand that has spilled outside the tray.

Color thresholding was applied after removing the highlights to isolate specific colors in the image (Figure 7). In this study, the HSV format was used for color thresholding. The HSV color space is more aligned with human visual perception than the RGB color space because it effectively separates intensity from hue (H) [19]. However, due to slight variations in lighting conditions within the experimental environment, there were disparities in color between the images, leading to suboptimal results in color thresholding as depicted in Figure 7(b2). Upon careful examination of sand characteristics, it was discovered that the sand could still be successfully distinguished and separated from the background.

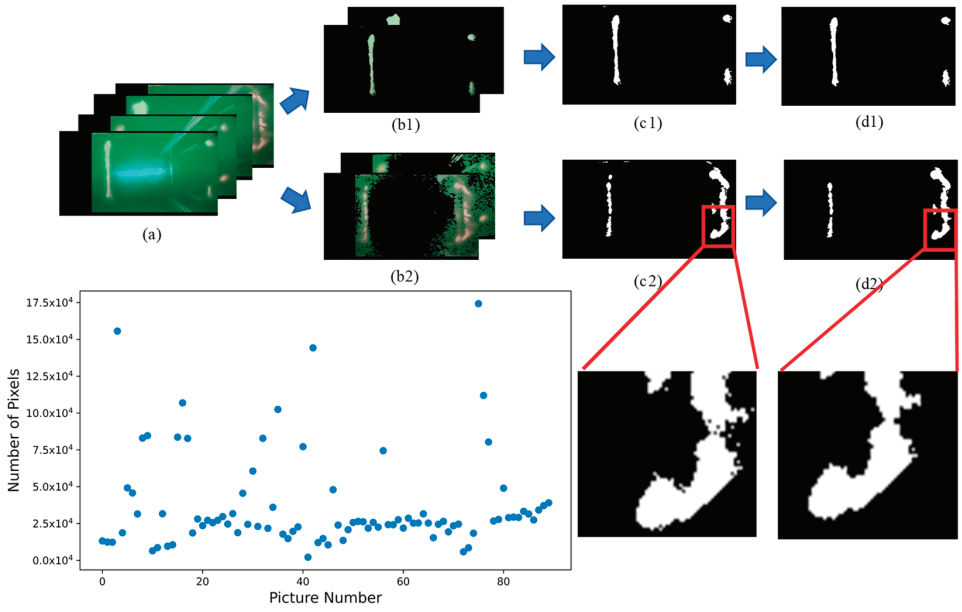


Figure 7. Sand area segmentation from the conveyor: (**a**) Image after preliminary preprocessing; (**b1**,**b2**) Results after the first color thresholding; (**c1**,**c2**) Binary images are generated after color thresholding with different segment methods; and (**d1**,**d2**) Images after noise reduction.

In contrast, the poorly separated areas occupy a relatively large area, as shown in Figure 7(b2). Therefore, the pixel values of the non-masked areas in all segmented images were counted. The area of all non-masked regions in the pictures was binary classified using K-means [20] to discriminate between the two segmentation situations (Figure 7(b1,b2)). The remaining 14 images were labeled low quantity and segmented using the YCbCr color space to produce binary images (Figure 7(c2)). A total of 76 images were directly converted into binary images (Figure 7(c1)), while the remaining 14 images were segmented directly into binary images (Figure 7(c2)). The reason for selecting this color space for segmentation is that the images clearly show uneven illumination, and the YCbCr color space is more suitable for image segmentation based on chromaticity and intensity [21]. Two sets of binary images with improved segmentation results were obtained after each set of images was processed separately, and the parameters were adjusted through continuous trial and error (Figure 7(c1,c2)). The following steps were taken to remove noise from the segmented binary images: (1) filling the 'holes' in the white areas using the built-in function in MATLAB® to fill the binary image and (2) setting all regions in the image with fewer than 500 pixels to black. Finally, the area of sand that leaked outside the tray was obtained

(Figure 7(d1,d2)) after these steps were taken. The area of each pixel in the processed images was determined during the preprocessing steps described in previous sections. Counting the number N of images with white pixels (with a value of 1) is essential to determine the area s of sand that has leaked outside the tray and passed (Equation (2)).

$$S = 0.19 \cdot N \tag{2}$$

To generate the depth information table (referred to as BWD) with a mask, the binary image (BW) representing the sand that spilled outside the tray after processing needs to undergo a dot product with the depth information table (D). Additionally, it is crucial to calculate the distance between the depth camera and the conveyor belt. Due to slight vibrations in the movement of the conveyor belt and the equipment during each experiment, there may be minor variations in the distance between the depth camera and the conveyor belt. In this process, the average depth information of the black region in the center of the binary image, measuring 100×100 pixels, is considered the distance between the depth camera and the conveyor belt. By subtracting the height (h) from the non-zero elements of the matrix (BWD), a corresponding height matrix (H) is obtained for each pixel block. Finally, the volume (V) of the sand that leaked outside the tray can be calculated using Equation (3).

$$H \odot S = V \tag{3}$$

2.7.3. Sand Surface Flatness

To calculate the flatness of the sand, it is necessary to localize the region of the sand. However, localizing this region can be a challenging task, mainly due to the mirror effect caused by the metal tray wall. This effect creates noise that closely resembles the appearance of the sand itself. In order to overcome this challenge, this study employed a combination of edge detection and color thresholding techniques. This approach enabled the segmentation of the sand region located in the middle of the tray with greater accuracy and precision.

First, RGB images (Figure 8a) are segmented using color thresholding. It is visible that the plate and the sand are brighter than other areas. This is because the HSV color space is more closely related to human perception of color than the RGB color space and can effectively distinguish the brightness and saturation levels of colors. Therefore, this study performed segmentation in the HSV color space and generated a mask image. After masking out the non-target areas at the edges accordingly, the plate's and sand's overall region can be obtained. The built-in function in MATLAB® was used to fill the binary image and remove the regions with fewer than 500 pixels (Figure 8c).

The RGB image is first subjected to a Gaussian filter using a 7×5 convolution filter, and the blurred image is then processed using the Canny operator for edge detection. Two different thresholds were set to detect edges in the image. The first threshold (set to 5) detects obvious edges, while the second threshold (set to 50) connects these edges. The convolution kernel is iteratively selected by trial and error to achieve the best results. This operation is performed three times using a 9×9 convolution kernel, each iteration using the same kernel size. After processing, a final binary image, the edge map, is obtained (Figure 8d). However, edge discontinuity is a common phenomenon in edge detection algorithms, and the Canny algorithm does not perform well in noise. After obtaining the edge map, this experiment also performed dilation and erosion operations. The dilation operation mainly increases the area of pixels with a value of 1 (white pixels). In contrast, the erosion operation mainly increases the size of pixels with a value of 0 (black pixels).

- By observing the image of the results, it can be seen that the unconnected parts of the edges are mainly concentrated in the lower part and the four corners. Therefore, it requires appropriately modifying the convolution kernel of the dilation operation to link the edges. After testing, a convolution kernel with a size of 9×7 was selected. This convolution kernel can dilate more in the horizontal direction and less in the

vertical direction while maintaining connectedness among all horizontal lower edges and achieving the highest possible boundary accuracy.

- To ensure the edge is as precise as possible, it is necessary to perform an erosion procedure after the dilation operation to reduce its expansion. After extensive testing, a convolution kernel with a size of 7×6 was eventually chosen for the erosion procedure. The reason for selecting an asymmetric convolution kernel with unequal sides is to correspondingly reduce the width of the edge caused by the dilation and to follow the dilation convolution kernel after trial and error (Figure 8e). Next, the binary image of the overall region of the plate and sand mentioned above is added using matrix addition. The image segmentation detects areas where pixel values are not equal to 1 and sets them to 0. The corresponding area of sand in the tray can be obtained by directly finding the largest connected region in the extracted image (Figure 8f). Then, denoising and edge smoothing operations are performed to obtain the image with the mask, as shown in (Figure 8g).
- In the last step, the sand region's binary image can be obtained. Then, the corresponding depth information of the sand in the tray can be obtained by taking the dot product of the binary image and depth point cloud information (Equation (4)). After denoising the point cloud information, the standard deviation is calculated to determine the flatness of the sand in the tray from the corresponding image.

$$I \odot D = SD \tag{4}$$

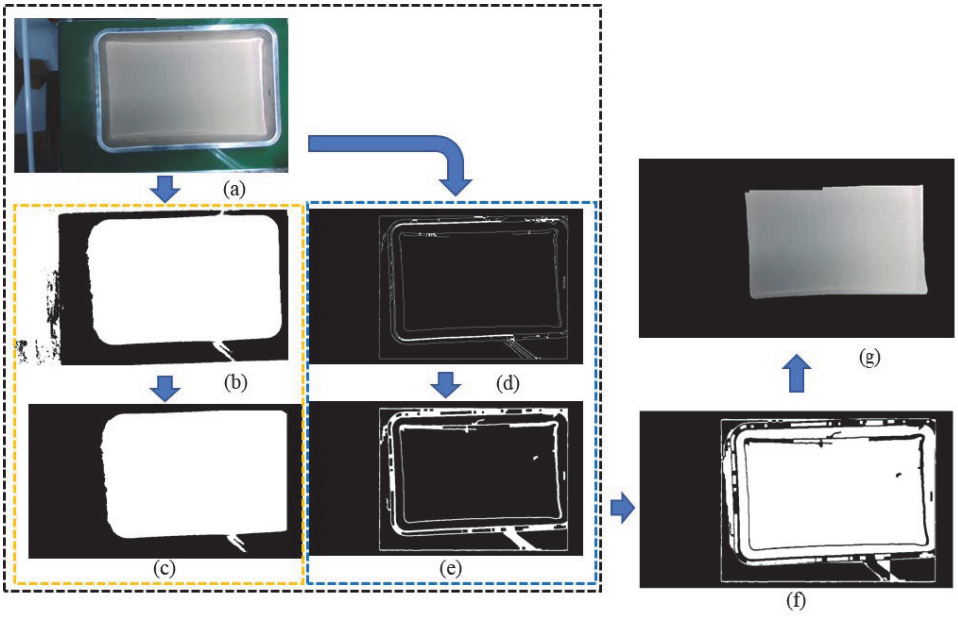


Figure 8. Flatness process: (**a**) RGB image; (**b**) Image after color threshold segmentation; (**c**) Image with masked edges; (**d**) Edge detection; (**e**) Dilation and erosion; (**f**) Image obtained by performing bitwise OR operation on two images; and (**g**) Target region.

2.8. Determining the Satisfactory Factor Combination

This experiment used the analytic hierarchy process (AHP) to discover a good combination of the three factors (i.e., sand weight in the tray, sand volume left on the conveyor, and sand surface flatness). AHP is a technique that may prioritize decisions based on pairwise comparisons and expert assessment [22]. The primary process of using AHP

is as follows: (1) data normalization; (2) establishing a judgment matrix; (3) consistency analysis; (4) calculating weights; and (5) calculating scores for each factor level combination. Normalization must be conducted as the first step to account for the significant difference in data magnitude between the three evaluation indicators. However, as weight serves as a positive indicator (where higher values indicate better performance), it is necessary to perform positive indicator normalization using Equation (5). Conversely, as volume and standard deviation act as negative indicators (where lower values indicate better performance), a negative indicator normalization using Equation (6) is required. Finally, the normalized values of the three indicators (weight, volume, and flatness) are obtained.

$$X_{normal} = \frac{X - X_{\min}}{X_{\max} - X_{\min}} \tag{5}$$

$$X_{normal} = \frac{X_{\max} - X}{X_{\max} - X_{\min}} \tag{6}$$

A judgment matrix was created using Saaty's scoring method to compare each pair of indicators, and the corresponding scoring criteria are presented in Table 1. Saaty's scoring criteria offer a clear and accurate evaluation of the level of different factors, and they are widely adopted because of their ease of comprehension. Based on the scores given by three experts and taking the median value, it was determined that the importance of flatness to weight and volume are 7 and 4, respectively. On a scale of 1 to 9, the importance of volume to weight is rated at 3, while flatness is rated at 1/4. Also, the importance of weight to volume and flatness are rated 1/3 and 1/7, respectively. The judgment matrix can be constructed as follows.

$$A = \begin{bmatrix} 1 & 0.33 & 0.14 \\ 3 & 1 & 0.25 \\ 7 & 4 & 1 \end{bmatrix} \tag{7}$$

Table 1. Saaty's scoring scale and its explanation.

Score	Explanation
1	Equal importance
3	Moderate importance
5	Very strong importance
7	Intermediate between extreme and very strong importance
9	Intermediate between very strong and strong importance
2, 4, 6, 8	The importance is somewhere in between

The maximum eigenvalue of the judgment matrix is $\lambda_{\max} = 3.032$. RI can be found to be 0.58, provided by Saaty's AHP method. Using Equation (8), the consistency index (CI) can be calculated using Equation (9), and the consistency ratio (CR) can be calculated.

$$CI = \frac{\lambda_{\max} - n}{n - 1} \tag{8}$$

$$CR = \frac{CI}{RI} \tag{9}$$

$$CR = 0.0311 < 0.1 \tag{10}$$

If CR is less than 0.1, it is considered to have passed the consistency check (Equation (10)) in this experiment. Then, the weights can be calculated using the arithmetic mean method, and the weights are shown in Table 2.

Table 2. Indicators and their corresponding weights.

Indicators	Weights
Weight	0.086
Volume	0.213
Flatness	0.701

By utilizing the weights derived from the AHP method, Equation (11) can be employed to compute the scores for all possible combinations of conditions.

$$score = 0.086 \times W + 0.213 \times V + 0.701 \times F \quad (11)$$

The experiment repeated the same factor level combination three times to prevent data randomness. The average values of weight, volume, and flatness indicators from three repetitions were used for those indicators in that factor level combination. Table 3 shows the experimental factors and levels.

Table 3. Factors and levels of the experiment.

Level	Factor A: Gate Opening Size (mm)	Factor B: Sensor Location (mm)	Factor C: Conveyor Speed (mm/s)
1	7.2	40.0	13.0
2	9.0	70.0	51.0
3	10.8	90.0	65.0

3. Results and Discussion

3.1. Weight of Sand in the Tray

Figure 9 shows the weight of sand in the tray under different factor combinations. It can be seen that when Factor A (gate opening size) and Factor C (conveyor speed) remain constant, the change in Factor B does not result in a significant weight variation, which means that the distance of the sensor has little effect on the weight of the sand in the tray. We calculate the standard deviation of Factor B under the same conditions as Factor A and Factor C. The standard deviations, relative to the mean of Factor B, are generally lower than 15%. The size of the opening plays a crucial role, particularly when it reaches the maximum opening size (A3). At this point, there is a significant difference in the weight of sand present in the tray compared to the two other opening sizes (A1 and A2). Additionally, the speed of the conveyor belt has a notable effect on the sand in the tray. As the conveyor belt speed increases, the weight of the sand in the tray decreases.

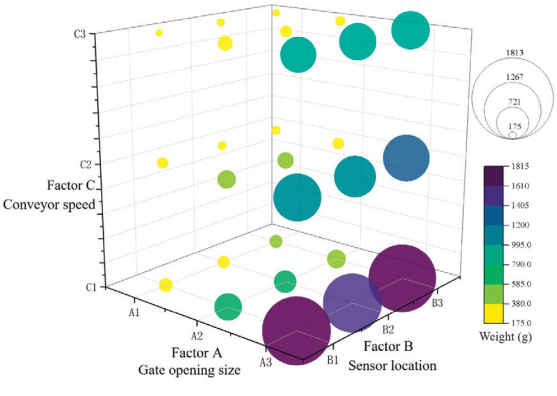


Figure 9. Weight data of sand in experimental trays for each group. (Factor A: Gate opening size; Factor B: Sensor location; and Factor C: Conveyor speed).

3.2. The Volume of Sand Outside the Tray

Figure 10 shows the volume of sand under different factor combinations. The volume of sand outside the tray (left on the conveyor) increases when the conveyor belt speed increases. When the conveyor belt speed is fast, the tray moves quickly and may exit the gate area before the gate is fully closed, causing sand to be left on the belt. Furthermore, the positioning of the sensor is also a significant factor affecting the performance. Placing the sensor closer to the device greatly reduces the volume of sand leaked outside the tray, while placing it farther away results in more sand leakage. When the gate opened before trays are delivered under the device, a large amount of sand will fall on the conveyor. If the device aperture is tiny and the sensor is located near the device, a large amount of sand will be wasted. This can occur even when the conveyor belt speed remains constant. On the other hand, the amount of sand lost is less if the device hole is wide and the sensor is positioned further away.

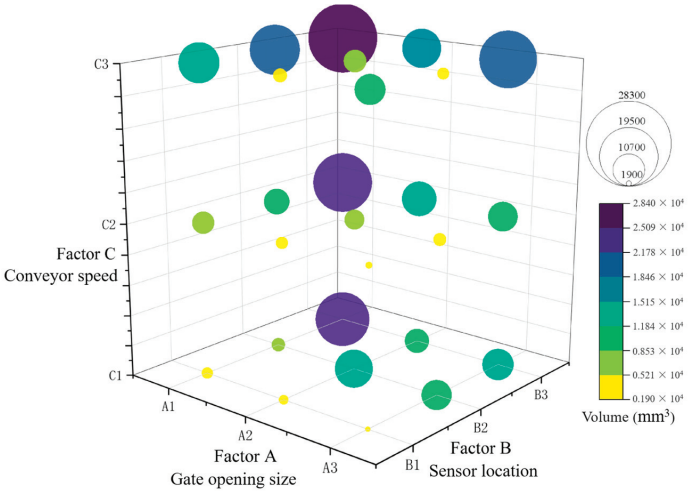


Figure 10. Volume data of sand leaked outside the tray for each group of experiments. (Factor A: Opening Size; Factor B: Sensor Location; and Factor C: Conveyor Speed).

3.3. Flatness of Sand Surface in the Tray

Figure 11 illustrates the variation in flatness across different factor combinations. The flatness of the sand within the tray is observed to be poor in cases where the opening size is either the smallest or the largest. This may be attributed to the uneven distribution of sand when the opening is small, and a higher quantity of sand falling simultaneously when the opening is wide. Moreover, the contact force between sand particles and the subtle vibrations resulting from the motor movement can also influence the final flatness of the sand within the tray. The experimental data indicate that, on average, the flatness level is better when using a moderate opening size (A2) compared to the other two sizes. Additionally, it is observed that a slower conveyor belt speed can yield a relatively flat sand surface when the opening is small. Conversely, a faster conveyor belt speed is required when dealing with a larger opening size to achieve a relatively flat sand surface in the tray. When the sensor is placed closer, the flatness of the sand in the tray tends to decrease. Conversely, when the sensor is positioned farther away, the flatness improves. This occurs because when the sensor is in close proximity to the sand outlet, the tray enters beneath the outlet before the device reaches the desired size, resulting in uneven sand distribution in front of the tray.

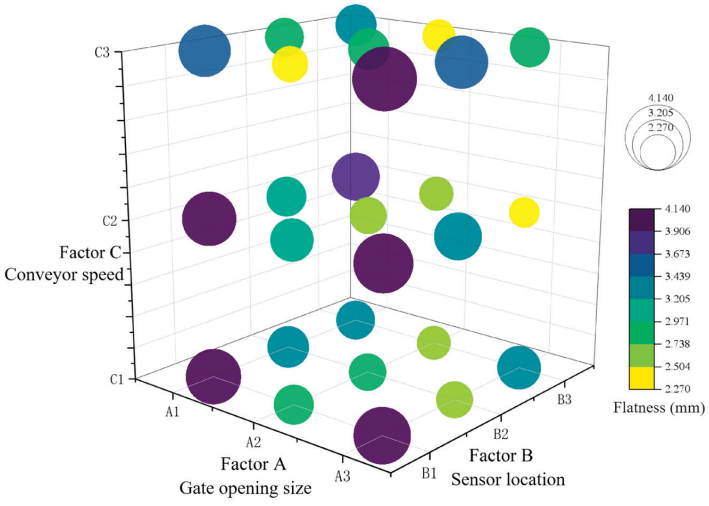


Figure 11. Data of sand flatness for each group of experiments (Factor A: Opening Size; Factor B: Sensor Location; and Factor C: Conveyor Speed).

3.4. Satisfactory Combination

Two combinations, A3B3C2 and A2B1C3, both yield scores above 0.8 when calculated using Equation (11). The combination A3B3C2 corresponds to an opening size of 10.8 mm, a sensor distance of 9.0 cm, and a conveyor belt speed of 5.1 cm/s. Similarly, the combination A2B1C3 represents an opening size of 9.0 mm, a sensor distance of 4 cm, and a conveyor belt speed of 6.5 cm/s. Various score values were tested based on the expert-assigned importance rules (where flatness is deemed more important than volume, and volume is considered more critical than weight). However, even with different score values, the combinations A3B3C2 and A2B1C3 consistently achieved the highest scores. According to our scoring calculation method, this experiment identifies A3B3C2 as the best combination while acknowledging that A2B1C3 is also a favorable choice. During the experiment, it was noticed that the flatness of the sand surface is affected by the sand-filling rate and the speed at which the tray moves (conveyor speed). The sand-filling rate is controlled by the gate opening size. Thus, regarding the sand surface flatness, the gate opening size and the speed of the conveyor are the two factors requiring investigation. Sand that falls onto the belt, not in the tray, is wasted. Adjusting the timing of the device's opening and closing can help reduce the amount of wasted sand. Since the time adjustment is related to the location of the sensor, the sensor installation place is another factor requiring investigation. Thus, the effects of three factors (i.e., gate opening size, conveyor speed, and sensor mounting location) on the machine performance should be studied.

4. Conclusions

This study focused on the development of an automated seeding machine, which involved four distinct phases: sand paving, seed positioning, watering, and sand covering. The performance of the machine was influenced by three factors: conveyor belt speed, gate opening size, and sensor horizontal distance from the sand gate. To evaluate its performance, three parameters were considered: the weight of sand in the tray, the volume of sand outside the tray, and the flatness of the sand. The analytic hierarchy process was employed to determine the optimal combination by exploring the relationship between these three assessment indicators.

The evaluation process placed significant emphasis on flatness due to its negative impact on the early germination rate detection of maize seeds. Considering the experiment's outcomes and utilizing the AHP approach, the following combination of factors is recom-

mended for achieving the most satisfactory performance: a gate opening size of 10.8 mm, a horizontal sensor distance of 9 cm from the gate, and a conveyor speed of 5.1 cm/s. This particular combination proves to be optimal for the automated seeding machine, ensuring the uniform distribution of a substantial volume of sand and maintaining desirable flatness.

The sand-spreading device examined in this experiment is a self-contained module with the potential for application in various related fields. This cost-effective and straightforward device offers the possibility of reducing expenses and enhancing profitability for companies. As a result, it holds promise for diverse industries that necessitate accurate and uniform sand or granular material distribution. Furthermore, the methods and evaluation apparatus employed in this study can be extrapolated to larger, more intricate sand-spreading machinery, offering valuable reference values for advancing precision agriculture's development.

Author Contributions: Conceptualization, B.F. and Z.Z.; methodology, Z.Z.; software, B.F.; validation, B.F., W.S. and Z.Z.; formal analysis, B.F.; investigation, B.F.; resources, Z.Z.; data curation, B.F.; writing—original draft preparation, B.F.; writing—review and editing, W.S. and Z.Z.; visualization, B.F.; supervision, Z.Z.; project administration, Z.Z.; funding acquisition, Z.Z. All authors have read and agreed to the published version of the manuscript.

Funding: This research is supported by the Chinese Universities Scientific Fund (funding no. 15053343).

Data Availability Statement: The data presented in this study are available on request from the corresponding author.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Ni, C.; Wang, D.; Vinson, R.; Holmes, M.; Tao, Y. Automatic inspection machine for maize kernels based on deep convolutional neural networks. *Biosyst. Eng.* **2019**, *178*, 131–144. [CrossRef]
2. Maize Production Share Worldwide by Country 2021/2022. Available online: https://www.statista.com/statistics/254294/distribution-of-global-maize-production-by-country-2012/ (accessed on 4 October 2022).
3. Li, C.; Raheja, A.; Still, D.W. Application of computer vision for lettuce seeds germination detection. In Proceedings of the International Conference on Image Processing, Computer Vision, and Pattern Recognition (IPCV), Las Vegas, NV, USA, 18–21 July 2011; p. 1.
4. Ducournau, S.; Feutry, A.; Plainchault, P.; Revollon, P.; Vigouroux, B.; Wagner, M.-H. Using computer vision to monitor germination time course of sunflower (*Helianthus annuus* L.) seeds. *Seed Sci. Technol.* **2005**, *33*, 329–340. [CrossRef]
5. Valiente-Gonzalez, J.M.; Andreu-Garcia, G.; Potter, P.; Rodas-Jorda, A. Automatic maize (*Zea mays*) kernel inspection system using novelty detection based on principal component analysis. *Biosyst. Eng.* **2014**, *117*, 94–103. [CrossRef]
6. Wittsell, L.E. *Effects of Soil Compaction on Plant Growth*; Kansas State University: Manhattan, KS, USA, 1964.
7. Wang, Y.; Yang, S.; Sun, J.; Liu, Z.; He, X.; Qiao, J. Effects of Tillage and Sowing Methods on Soil Physical Properties and Corn Plant Characters. *Agriculture* **2023**, *13*, 600. [CrossRef]
8. Roquelaure, Y.; Jégo, S.; Geoffroy-Perez, B.; Chazelle, E.; Descatha, A.; Evanoff, B.; Garlantézec, R.; Bodin, J. Carpal tunnel syndrome among male French farmers and agricultural workers: Is it only associated with physical exposure? *Saf. Health Work* **2020**, *11*, 33–40. [CrossRef] [PubMed]
9. Pankaj, G.; Kadiwala, M. Comparative study of different soil covering and pressing devices. *Trends Biosci.* **2017**, *10*, 2452–2455.
10. Xu, H.; Tian, D.; Liu, J.; Niu, Z.; Li, Q. Operation analysis and parameter optimization of drum type soil-covering device. *Int. J. Agric. Biol. Eng.* **2021**, *14*, 123–129. [CrossRef]
11. Zollhöfer, M.; Stotko, P.; Görlitz, A.; Theobalt, C.; Nießner, M.; Klein, R.; Kolb, A. State of the art on 3D reconstruction with RGB-D cameras. *Comput. Graph. Forum* **2018**, *37*, 625–652. [CrossRef]
12. Zhou, T.; Fan, D.-P.; Cheng, M.-M.; Shen, J.; Shao, L. RGB-D salient object detection: A survey. *Comput. Vis. Media* **2021**, *7*, 37–69. [CrossRef] [PubMed]
13. Vit, A.; Shani, G. Comparing rgb-d sensors for close range outdoor agricultural phenotyping. *Sensors* **2018**, *18*, 4413. [CrossRef] [PubMed]
14. Xia, C.; Wang, L.; Chung, B.-K.; Lee, J.-M. In situ 3D segmentation of individual plant leaves using a RGB-D camera for agricultural automation. *Sensors* **2015**, *15*, 20463–20479. [CrossRef] [PubMed]
15. Skoczeń, M.; Ochman, M.; Spyra, K.; Nikodem, M.; Krata, D.; Panek, M.; Pawłowski, A. Obstacle detection system for agricultural mobile robot application using RGB-D cameras. *Sensors* **2021**, *21*, 5292. [CrossRef] [PubMed]

16. Pollefeys, M.; Koch, R.; Gool, L.V. Self-calibration and metric reconstruction inspite of varying and unknown intrinsic camera parameters. *Int. J. Comput. Vis.* **1999**, *32*, 7–25. [CrossRef]
17. Telea, A. An image inpainting technique based on the fast marching method. *J. Graph. Tools* **2004**, *9*, 23–34. [CrossRef]
18. Oagile, O.; Gabolemogwe, P.; Matsuane, C.; Mathowa, T. Effect of container size on the growth and development of tomato seedlings. *Int. J. Curr. Microbiol. Appl. Sci.* **2016**, *5*, 890–896. [CrossRef]
19. Sural, S.; Qian, G.; Pramanik, S. Segmentation and histogram generation using the HSV color space for image retrieval. In Proceedings of the International Conference on Image Processing, Rochester, NY, USA, 22–25 September 2002.
20. Bock, H.-H. Clustering methods: A history of k-means algorithms. In *Selected Contributions in Data Analysis and Classification*; Springer: Berlin/Heidelberg, Germany, 2007; pp. 161–172.
21. Shaik, K.B.; Ganesan, P.; Kalist, V.; Sathish, B.; Jenitha, J.M.M. Comparative study of skin color detection and segmentation in HSV and YCbCr color space. *Procedia Comput. Sci.* **2015**, *57*, 41–48. [CrossRef]
22. Saaty, T.L. Decision making with the analytic hierarchy process. *Int. J. Serv. Sci.* **2008**, *1*, 83–98. [CrossRef]

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.

agriculture

MDPI

Review

Research Progress on Autonomous Operation Technology for Agricultural Equipment in Large Fields

Wenbo Wei [1], Maohua Xiao [1,*], Weiwei Duan [1], Hui Wang [2], Yejun Zhu [1], Cheng Zhai [2] and Guosheng Geng [1]

[1] College of Engineering, Nanjing Agricultural University, Nanjing 210031, China; wboclear@stu.njau.edu.cn (W.W.); 2024112031@stu.njau.edu.cn (W.D.); yjzhu@njau.edu.cn (Y.Z.); gsgeng@njau.edu.cn (G.G.)

[2] Weichai Lovol Intelligent Agricultural Technology Co., Ltd., Weifang 261206, China; wanghui4@lovol.com (H.W.); zhaicheng@lovol.com (C.Z.)

* Correspondence: xiaomaohua@njau.edu.cn; Tel.: +86-13951756153

Abstract: Agriculture is a labor-intensive industry. However, with the demographic shift toward an aging population, agriculture is increasingly confronted with a labor shortage. The technology for autonomous operation of agricultural equipment in large fields can improve productivity and reduce labor intensity, which can help alleviate the impact of population aging on agriculture. Nevertheless, significant challenges persist in the practical application of this technology, particularly concerning adaptability, operational precision, and efficiency. This review seeks to systematically explore the advancements in unmanned agricultural operations, with a focus on onboard environmental sensing, full-coverage path planning, and autonomous operational control technologies. Additionally, this review discusses the challenges and future directions of key technologies for the autonomous operation of agricultural equipment in large fields. This review aspires to serve as a foundational reference for the development of autonomous operation technologies for large-scale agricultural equipment.

Keywords: agricultural equipment; autonomous operation; environmental sensing; complete-coverage path planning

Citation: Wei, W.; Xiao, M.; Duan, W.; Wang, H.; Zhu, Y.; Zhai, C.; Geng, G. Research Progress on Autonomous Operation Technology for Agricultural Equipment in Large Fields. *Agriculture* **2024**, *14*, 1473. https://doi.org/10.3390/agriculture14091473

Academic Editor: Lixia Hou

Received: 16 July 2024
Revised: 19 August 2024
Accepted: 27 August 2024
Published: 29 August 2024



Copyright: © 2024 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (https://creativecommons.org/licenses/by/4.0/).

1. Introduction

Agriculture is a typical labor-intensive industry that relies heavily on labor [1]. With the aging of the population, the agriculture in countries where smallholder farming is the prevalent form of agriculture has been severely affected. These effects are more severe in countries such as China, where smallholder farmers predominate. Compared to 1990, about 4 million hectares of cropland were abandoned in China in 2019 [2]. In the future, the aging population in China is expected to accelerate further, necessitating urgent measures to reduce agriculture's heavy reliance on manual labor [3–5]. To cope with the status quo of smallholder agriculture, new agricultural models have been encouraged by the Chinese government, mainly including family farms, new rural cooperatives, and large-scale farms. The new model is conducive to promoting modern agriculture towards an intelligent and scaled agricultural production model [6,7]. The technology of autonomous operation of agricultural equipment in large fields can be adapted to new modes of agricultural production, which is an important method to improve productivity and reduce labor intensity, and it helps to alleviate the impact of population aging on agricultural production [8–13]. The key technologies for autonomous operation of large-scale agricultural equipment mainly include onboard environmental sensing technology, complete-coverage path-planning technology and autonomous operation control technology, as shown in Figure 1.

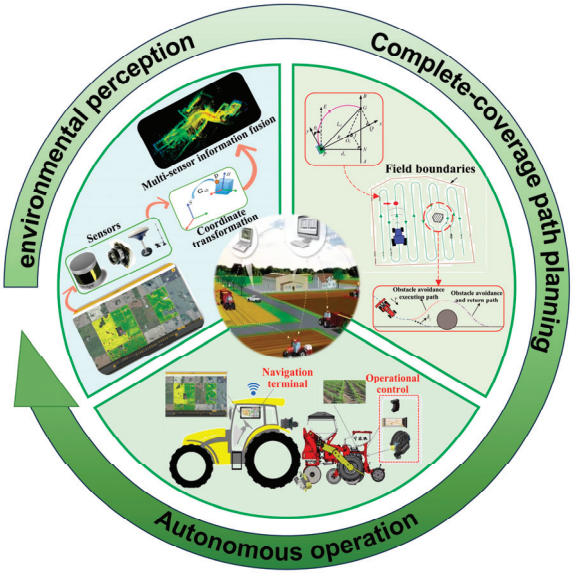


Figure 1. Schematic diagram of the key technology for autonomous operation of agricultural equipment in large fields. This figure illustrates the comprehensive workflow, from multi-sensor environmental data acquisition to complete-coverage path planning, culminating in the autonomous operation of agricultural machinery. Collectively, these factors influence the overall effectiveness of autonomous operations in large-scale agricultural machinery.

Onboard environmental sensing technology is the primary condition for agricultural machines to achieve unmanned operations [14], and it is also the key to ensuring the safe and efficient operation of unmanned agricultural machines in complex and variable unstructured farmland environments. Despite the remarkable advancements achieved in unmanned farm machinery, its practical application continues to confront numerous challenges, stemming from the intricacies and variability of the farmland environment. For example, a single sensor is often utilized to gather environmental information during unmanned operation processes such as cultivation, management, and harvesting [15]. However, a single sensor can only obtain one-sided information when sensing the environment, and there are problems, such as insufficient autonomous sensing ability and poor environmental adaptability. Agricultural machines equipped with multiple environment-aware sensors have improved stability and reliability in unmanned operations, but the application of multi-sensor information fusion algorithms still faces challenges [16]. Hence, analyzing the data characteristics of sensors employed in environmental sensing technologies, while fully leveraging the redundancy and complementary aspects of multi-sensor data, is crucial for developing stable and reliable environmental sensing systems.

The key to realizing the unmanned operation of agricultural equipment lies in the complete-coverage path-planning technology. For different operational tasks and farmland environments, the goal is often to minimize operational costs and cover a wide range of operations [17]. Path-planning technology is employed to devise the optimal driving path for achieving efficient and precise farmland operations. Path-planning techniques are typically categorized into global path planning and local path planning [18,19]. Global path planning aims to plan the optimal path in the whole operation area to maximize the coverage and minimize the costs. Local path planning focuses on responding to real-time changes in the environment, ensuring that agricultural machinery can safely and efficiently accomplish operational tasks in a dynamic environment. At this stage, the main problems faced by path-planning technology include operational efficiency, operational plot coverage, operational safety, and other aspects. Despite significant advancements in

path-planning technology, it still faces numerous challenges and requires further research and improvement. Enhancing its adaptability to complex environments, real-time responsiveness, and multi-objective optimization is essential to meet the demands of modern agricultural production.

Autonomous operation control technology is the core for enabling the unmanned operation of agricultural equipment, encompassing two key aspects: motion control and operational control [20–22]. Motion control primarily enhances path-tracking accuracy through speed control and steering control. Operation control addresses the specific tasks performed by agricultural machinery in various operations, such as seeding, fertilizing, spraying, and harvesting [23,24]. Motion control and operation control technologies complement each other to ensure that intelligent farm machinery can efficiently and accurately complete operational tasks in the complex and changing farmland environment. In farmland environments, where there are errors in existence due to factors such as undulating terrain and changes in soil type, it is frequently challenging for motion control methodologies to accurately adhere to a predetermined path. Moreover, operational control is essential for enhancing the autonomous operational capabilities of agricultural machinery, achieving fully unmanned operations, and advancing agriculture towards intelligence and precision. Although some common control methods can mitigate motion control errors and improve the intelligence of operational control, there is still a need for some better methods to improve motion control and operational control effectiveness.

In conclusion, onboard environmental sensing technology, complete-coverage path-planning technology, and autonomous operation control technology are the key technologies for the unmanned operation of agricultural equipment in large fields. This paper focuses on analyzing the current research status of these three components. Currently, numerous scholars have conducted extensive studies on these three components. However, intelligent agricultural machines, designed for unmanned operations in large fields, are rapidly advancing, but progress overall is uneven. This uneven progress hinders scholars from accurately pinpointing the direction of key technology development in this field. Therefore, it is crucial to systematically summarize the research progress on the key technology of unmanned field operations for large agricultural equipment. The main objective of this paper is to review and analyze representative literature on the key technology of the unmanned field operation of large-scale agricultural equipment, providing a reference for the development of unmanned field operation technology for large agricultural machinery.

2. Onboard Environmental Sensing Technology

Onboard environmental sensing technology is primarily employed to address the complexities and dynamic nature of the operational environment by utilizing localization systems and sensory devices mounted on agricultural machinery. The operational state of the machinery is then continuously adjusted in real time, informed by the data on the terrain, obstacles, and other environmental factors [25,26]. Subsequently, the system swiftly modifies the machine's operational status in response to real-time data inputs. Compared to onboard sensing tasks for road vehicles, farmland lacks distinctive structured features, making it essential to create environment models of typical farmland elements using sensor data for unmanned farm machinary operations. Key sensory elements in farmland comprise information on farmland boundaries, crop rows, and obstacles within the farmland (as illustrated in Figure 2). Currently, the onboard environmental sensing tasks of intelligent agricultural equipment in unstructured farmland environments are mainly focused on localization data acquisition, farmland boundary detection, navigation line extraction, and obstacle detection [27–30]. The main methods are vision and radar detection, which are committed to improving the robustness of the sensing system and enhancing the autopilot and autonomous operation performance of the agricultural equipment.

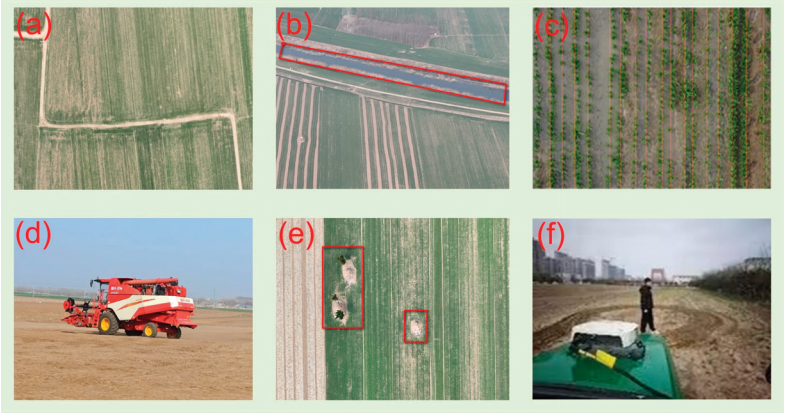


Figure 2. Typical perceptual elements in farmland: (**a**) farm boundary; (**b**) ditch; (**c**) crop row; (**d**) operating farm machinery; (**e**) stationary obstacle; (**f**) person.

The acquisition of precise positioning data is the fundamental prerequisite for enabling the navigation operations of agricultural machinery. The Global Navigation Satellite System (GNSS) has emerged as a universally recognized and widely adopted technology for tracking and positioning agricultural machinery, providing crucial data, such as heading, speed, and time [31]. The GNSS encompasses multiple satellite positioning systems, including GPS, BDS, GLONASS, and Galileo, rather than referring to a singular satellite system [32,33]. While the GNSS can achieve positioning accuracy within a range of a few meters to several meters in open environments, the autonomous navigation of agricultural machinery demands centimeter-level precision, rendering typical GNSS accuracy insufficient for such applications [34–36]. Real-Time Kinematic (RTK) is a differential technique, used for the real-time processing of carrier-phase observations from two measurement stations, encompassing both traditional and network RTK [37,38]. The RTK system, an enhancement technology built upon standard GNSS, significantly improves positioning accuracy from meters to centimeters via differential algorithms, thereby offering essential spatial data support for the implementation of precision agriculture [39]. In recent years, extensive research efforts have been undertaken by scholars to advance navigation and localization techniques grounded in GNSS technology. Yue et al. [40] designed an automatic navigation system for a tracked orchard sprayer based on GNSS technology. The navigation operation accuracy was improved by compensating the effect of machine vibration on the localization data. Kaivosoja et al. [41] developed a GNSS error simulator to simulate the positioning error and positioning reliability of GNSS in various typical agricultural scenarios. Lee et al. [42] designed a dual antenna to receive localization information so that the harvester can obtain the current direction even under stationary working conditions. Although RTK-GNSS achieves high positioning accuracy in open environments, farmland conditions are often complex and variable, with frequent challenges, such as signal occlusion, multipath effects, and environmental obstacles, which can compromise positioning accuracy [43–45]. Consequently, multi-sensor information fusion technology is being progressively integrated into agricultural machinery navigation, harnessing complementary advantages to enhance overall localization performance. Jing et al. [46] designed an automatic navigation grading system for GNSS/INS, which utilized adaptive square root Kalman filtering to fuse the data from GNSS and INS to improve the operational efficiency of grading machines. Li et al. [47] fused GNSS positioning equipment and an inertial measurement unit based on a fuzzy adaptive finite impulse response Kalman filter algorithm to enable high positioning accuracy and the stability of agricultural implements. Xu et al. [48] proposed a backpropagation neural network-based GNSS/INS/OD/NHC adaptive combined navigation method considering vehicle motion, which fully considered

the relationship between forward speed, heading angular velocity, and lateral velocity. A review of the literature reveals that multi-sensor fusion localization can leverage data from additional sensors to ensure system stability and continuity in the event of sensor failure. However, practical applications of multi-sensor fusion localization continue to face challenges, including the complexity of data fusion algorithms and time synchronization issues among sensors.

Field boundary detection and navigation line extraction are important foundations for the autonomous operation capability of agricultural machinery. Currently, the methods of farmland boundary detection by onboard sensing systems of agricultural machines are mainly based on sensors such as mono/binocular vision and LiDAR. Hong et al. [48] designed a field boundary recognition and ranging method based on a binocular vision approach. Multi-feature fusion and split-tree algorithms were applied for cost calculation and cost aggregation respectively. Using the multi-scale cost aggregation framework, the linkage of matching costs at different scales was established to enhance the reliability in weak texture regions. Then, the interference elimination method based on point-cloud continuity was designed according to land cultivation and crop growth. Adaptive extraction for point-cloud boundary is realized, and accurate ridge boundary lines are identified, which improves the stereo-matching accuracy and speed of farmland images. Wang et al. [49] designed the detection method of non-regular headland boundary lines based on machine vision technology; this method provides a baseline for the automatic navigation and turning of agricultural equipment at the boundary of the farmland to achieve the whole autonomous navigation operation of agricultural machinery. Li et al. [50] proposed a machine vision-based headland boundary detection method for paddy fields, which introduced depth information, in addition to RGB images, to improve its detection accuracy. A deep learning network combining a convolutional neural network and a recurrent neural network was constructed for headland semantic segmentation. Meanwhile, an interactive attention module is proposed to fuse the complementary information in RGB-D images adaptively. In addition, the image preprocessing technique and the proposed distance-based clustering algorithm for boundary points are applied to the headland segmentation mask to obtain the boundary line of the farmland. This process provides technical support for the automatic turning strategy of the agricultural machine. There often exists some highly non-uniform vegetation at the field boundaries [51], which leads to difficulties in extracting the field boundaries. Although the point-cloud data acquired by LiDAR at the field boundary has irregular height characteristics, the data still have certain linear characteristics [52–54]. Therefore, strengthening the linear features of the point-cloud data becomes an important method for improving the accuracy of boundary extraction by LiDAR sensors.

Navigational line extraction is also crucial for improving the autonomous operation of agricultural machinery. Hou et al. [55] proposed a lightweight path recognition model based on the U-Net model, which solves the problem of poor accuracy and real-time performance of the method for recognizing navigation paths for fruits between ridges. The model can also improve the real-time navigation and accuracy requirements. Gong et al. [56] effectively extracted navigation lines based on the composite localization points of corn and stems and solved the problem regarding the difficult extraction of navigation lines in corn fields with weed interference. Diao et al. [57] proposed an algorithm for the extraction of navigation lines of corn-spraying robots based on an improved YOLOv8s network. The algorithm addressed the low accuracy of navigation line extraction for corn-spraying robots in a complex farmland environment and improved the accuracy of the autopilot feature. In addition, LiDAR is an important sensor for navigation line extraction. For example, it can be used for crop row identification and navigation line extraction within regions of interest in the LiDAR scanning area using the least squares method [58,59]. In addition, Wang et al. [60] designed a navigation-line extraction method for an autonomous spraying machine in an orchard based on LiDAR. Liu et al. [61] et al. proposed a tree row recognition and navigation method based on the fusion of least squares and support vector machines.

Using the least squares fitting method, combined with the SVM algorithm, to identify tree rows, the centerline between rows in the orchard is predicted, which improves the accuracy of the autonomous navigation of orchard equipment for the orchard environment.

Currently, farmland boundary detection and navigation line extraction are mostly based on visual schemes, mainly image grayscale processing, image segmentation, navigation line extraction, navigation line tracking, and other steps. The focus is on solving the effect of natural illumination on the robustness of the detection system. The commonly used methods for farmland boundary and navigation line extraction research are shown in Table 1.

Table 1. Characterization and application of commonly used sensors for field boundary and navigation line extraction.

Method	Sensor Type	Characteristics	Sensing Task
Vision Sensors	Monocular Camera	Monocular cameras are low cost and provide rich image information, but lack depth data and are susceptible to environmental influences.	Farmland boundary detection, navigation line extraction
	Binocular Camera	Binocular cameras can provide rich image information and highly reliable depth information, but the configuration and calibration are more complicated; the computation is large, and parallax calculation depends on computing resources.	Farmland boundary detection, navigation line extraction
	RGB-D Camera	RGB-D camera can provide an RGB map and a depth map, and the calculation amount is small. However, the measurement range is narrow, the noise level is high, the field of view is small, and it is easily interfered with by daylight.	Farmland boundary detection
Radar Sensor	Lidar	LIDAR is highly accurate, stable, and reliable. However, it has a high cost, is susceptible to dust interference with the limited detection range, and cannot recognize color and texture in farmland boundary identification and navigation line extraction.	Navigation line extraction

Obstacle detection is an important method to improve the level of autonomous operation of farm machinery in large fields and to ensure safe operation. Currently, the research on field obstacle detection based on the onboard sensing system of agricultural machinery mainly includes methods based on vision detection, LiDAR detection, millimeter-wave radar detection, and multi-sensor fusion detection [62], as shown in Figure 3.

In terms of visual detection, Li et al. [63] proposed a method based on a deep neural network to detect obstacles such as people, concrete pillars, and utility poles in the orchard environment. Their method can provide a basis for intelligent orchard robots to detect obstacles automatically and distinguish traversable and non-traversable areas for obstacle avoidance. Furthermore, Xue et al. [64] improved the YOLOv5s algorithm based on the k-means clustering algorithm and CIoU loss function to improve the detection speed of obstacles in farmlands. The algorithm also has reliable detection accuracy for small-target obstacles. Xu et al. [65] detected moving obstacles in panoramic images based on panoramic camera vision and the Lucas–Kanade optical flow algorithm. This method has good transient and detection results for dynamic obstacle detection. Du et al. [66] introduced spatial attention and deformable convolution into the Mask R-CNN model to construct a segmentation algorithm for unstructured farmland obstacles. The model can control the detection speed while still maintaining a reliable model detection accuracy. Deep learning models perform well in image processing and feature extraction. Deep learning methods have become some of the most widely used techniques in visual obstacle detection. Representative research methods mainly include YOLO, Faster R-CNN, U-Net, and RetinaNet. Deep learning models can automatically learn and extract complex features in images, greatly improving the accuracy and robustness of obstacle detection.

The commonly used camera types and features when using deep learning methods for obstacle detection are shown in Table 2.

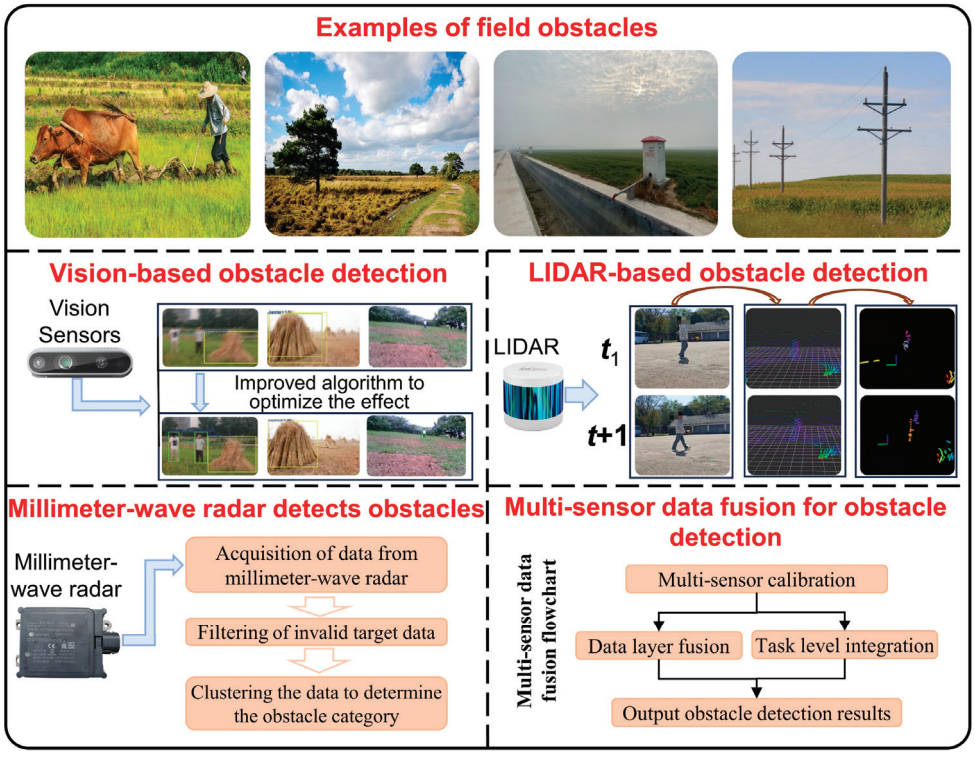


Figure 3. Methods for detecting obstacles in agricultural fields.

Table 2. Deep learning of camera types and characteristics commonly used in obstacle detection.

Camera Type	Features	Advantages	Common Cameras
RGB Cameras	A standard color camera that captures images in the red, green, and blue color channels.	RGB cameras provide rich color and texture information that helps distinguish between different types of obstacles, are low cost, and are easy to integrate and deploy.	Logitech C920, Sony Alpha Series (Logitech, Lausanne, Switzerland)
Depth Camera	In addition to capturing RGB images, it also acquires depth information for each pixel.	Combining depth information and RGB images improves the accuracy and reliability of obstacle detection, providing more precise obstacle localization, especially in complex environments.	Intel RealSense (Intel, Santa Clara, CA, USA), Microsoft Kinect 360 (Microsoft 360, Washington, DC, USA)
Stereo Camera	Captures stereo images through two cameras and uses parallax to calculate depth information.	Provides high-precision depth perception for fine obstacle detection tasks and is more reliable than a single depth camera in terms of detection accuracy and range.	ZED Series (ZED Series, San Francisco, America), Bumblebee2 (Teledyne FLIR, Washington, DC, USA)
Panoramic Camera	Capable of capturing images or videos with a 360-degree field of view.	In obstacle detection, it provides a comprehensive view of the environment, reduces blind spots, and improves the coverage and accuracy of obstacle detection.	Ricoh Theta (RICOH, Tōkyō, Japan), Insta360 Pro (insta360, Shenzhen, China)

In terms of obstacle detection by LiDAR and millimeter-wave radar, Shang et al. [67] proposed a 3D laser point-cloud method for detecting field obstacles based on Euclidean clustering. First, the voxel down-sampling method and RANSAC algorithm are utilized to

filter the point cloud and differentiate between ground and above-ground objects. Then, the obstacles are recognized based on the Euclidean clustering frame of the K-D tree, which achieves the detection of agricultural implements, haystacks, field ridges, short houses, and trees on both sides of the field road. T. Wang et al. [68] proposed an adaptive method for the real-time 3D detection of obstacles for a single sub-domain based on the semantic–geometric–intensity fusion strategy. The inaccurate 3D detection of obstacles in the absence of many samples is addressed, and the experimental results demonstrate the high accuracy and efficiency of the proposed method in performing obstacle detection. Using a millimeter-wave radar for farmland obstacle detection has the problem of high computation because of the presence of many invalid targets in the output data. Therefore, Xue et al. [69] addressed this issue by filtering out empty targets, pseudo-targets, and non-threatening data based on the invalid target filtering method. The above literature uses LiDAR or millimeter-wave radar as the detection sensor, which achieves better detection results. However, there is still the problem of high computation, which leads to low detection efficiency and is not easy to deploy.

Vision-based methods for obstacle detection are low cost and have advantages in measuring object height and width, as well as in recognition accuracy. They can also provide rich planar information. However, the role of vision distance and ranging accuracy in these methods are not as good as those in a millimeter-wave radar. Moreover, these methods are easily affected by light and weather [70]. LiDAR can obtain accurate position data while not being affected by light conditions. However, when a large amount of dust exists in the farmland environment, features such as shape and texture information of the obstacles are more difficult to obtain. Single sensors are still deficient in terms of reliability, robustness, and accuracy in obstacle detection. Multi-sensor information fusion methods for obstacle detection synthesize heterogeneous information from different sensors and have their respective advantages. Hence, applying these methods has become the main trend in obstacle detection. Lv et al. [71] designed a decision-level fusion algorithm by combining the advantages of millimeter-wave radar in range and speed measurement and the advantages of the camera in type recognition and lateral localization. After calibrating the outer parameters of the millimeter-wave radars and the inner and outer parameters of cameras, the obstacle detection test was conducted in a ROS environment. Comparative experiments with sensor fusion algorithms showed that the detection accuracy of the decision-level fusion algorithm was 95.19%, which is higher than that of feature-level and data-level fusion by 4.38% and 6.63%, respectively. Cai. [72] realized the temporal and spatial fusion of vision and millimeter-wave radars based on a fusion detection technology for a vision–millimeter-wave radar. The task of detecting obstacle dimensions in the visual depth map was accomplished using the effective target selected by the millimeter-wave radar as the seed point. The problem of detecting the spatial location and dimensional information of obstacles in front of agricultural machines was solved. Kragh et al. [73] used conditional random fields to fuse LiDAR and camera sensing probabilistically with semantic segmentation and combined appearance-based and geometric detection methods. Finally, a multimodal fusion algorithm was applied to achieve obstacle detection in agriculture with moving ground vehicles. Multi-sensor fusion sensing technology has shown great advantages in farmland obstacle detection and will be gradually applied to farmland obstacle detection in the future. The farmland environment is complex and variable, and a single sensor is prone to blind spots or insufficient information in obstacle detection. Multi-sensor fusion can comprehensively utilize the advantages of different sensors to make up for the shortcomings of a single sensor. The core of fusion technology is to solve the problem of the fusion of heterogeneous data information from different sensors. By fusing data from multiple sensors, such as RGB cameras, depth cameras, and LIDAR, more comprehensive and precise environmental information can be obtained to improve the accuracy and reliability of obstacle detection.

3. Complete-Coverage Path-Planning Technology

Based on farmland information, operation types and sensory data, complete-coverage path-planning technology plans optimal paths and commands the underlying modules to deal with complex farmland environments effectively. Currently, the task of farm machine path planning in unstructured farmland environments mainly focuses on global path planning and local path planning, with the goal of improving operational efficiency, operational farmland coverage, and operational safety to adapt to the complex and changing operational environments in different farmlands and fields [74]. The classification of common path-planning algorithms in path-planning technology and their range of applications is presented in Table 3. To perform specific agricultural tasks such as deep loosening, targeting, ridging, rotary ploughing, seeding, and applying pesticides, when operating unmanned vehicles in a field, large agricultural equipment often follows a path comprising straight lines and turns. Unlike standard roads, agricultural fields generally have irregular shapes (depicted in Figure 4). Achieving smooth path trajectories and velocity continuum is challenging when relying solely on a combination of straight lines and turns for complete-coverage path planning. Therefore, in recent years, researchers have proposed numerous algorithms to address complete-coverage path-planning problems in agricultural fields. In this chapter, the representative algorithms for complete-coverage path-planning technology are classified into two categories: classical path-planning algorithms and bionics-based path-planning algorithms.

Table 3. Classification and application scope of common path-planning algorithms.

Classification	Common Algorithms	Common Application Areas
Algorithms based on graph search	Dijkstra, A *, D *	Global path planning
Algorithm based on sampling	RRT	Global path planning
Algorithms based on artificial potential fields	Artificial potential field method	Local path planning
Algorithms based on curve fitting	Arcs and straight lines, polynomial curves, spline curves, Bessel curves, differential flatness	Local path planning
Algorithms based on numerical optimization	Describing and solving planning problems using objective functions and constraints	Local path planning
Intelligent algorithms based on bionics	Genetic algorithms, particle swarm optimization algorithms, ant colony algorithms	Global path planning, local path planning

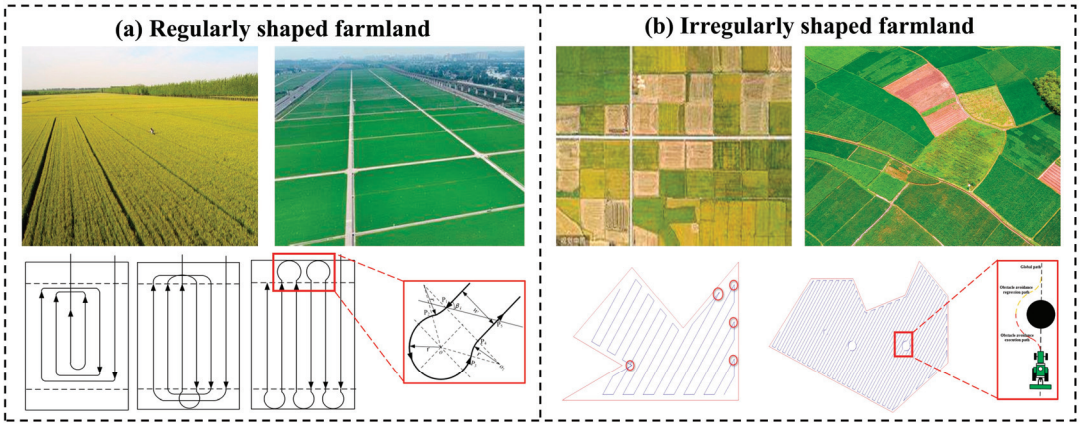


Figure 4. Shapes of typical regular and irregular fields.

3.1. Classical Path-Planning Algorithm

Global path planning is to plan an optimal operation path to carry out the operation by combining the positioning information after determining the operation plot. Local path planning integrates the real-time information obtained from the intelligent farm machinery vehicle-mounted sensing system, performs real-time path planning, and adjusts the operational status. The main classical path-planning algorithms currently applied are search-based global path-planning algorithms, sampling-based global path-planning algorithms, artificial potential field-based local path-planning algorithms, local path-planning algorithms based on curve fitting (arcs and straight lines, polynomial curves, spline curves, Bezier curves, and differential flatness), and numerical optimization-based local path-planning algorithms.

Shen et al. [75] proposed a primitive optimization strategy to correct the inter-row path points based on the tree row position information. This strategy is a response to the traditional A * algorithm's problem of path point offsets that are not suitable for direct substitution optimization. Xu et al. [76] designed an improved A * algorithm that reduces the number of turns in global path planning and improves path stationarity and operational efficiency. The RRT algorithm is slow to search when used for global path planning, and when the number of iterations is small, the feasibility of the planned global path is poor. Feng et al. [77] added global random sampling and key area sampling strategies to the RRT algorithm to improve the global search capability and the effect of optimal paths. Kong et al. [78] proposed an improved A * path-planning algorithm based on multi-constraint Bessel curves. The approach combines robot kinematic constraints with Bézier curves to smooth the turning paths of the A * algorithm, integrating multiple objectives, such as minimum turn radius and continuous curvature, to achieve global path planning. The literature [79,80] addresses the problem that traditional artificial potential field methods tend to fall into local minima and suffer from goal unreachability during path planning, by introducing an annealing algorithm and obstacle repulsive potential field function method to avoid the falling into local optimum and unreachable goal phenomena. Boryga et al. [81] proposed a path-planning method based on polynomial transition curves to plan the trajectory of a machine turning in a field by using polynomial transition curves to reduce the working and non-working distances and to improve the efficiency and energy utilization of the machine. In addition to the above methods, the method of improving the local path-planning algorithm by utilizing the circular arc tangent law has been well applied to tracked harvesters (shown in Figure 5). The improvement reduces the frequency of steering control of tracked harvesters, which in turn improves driving stability and operational efficiency [82–86]. There is no essential difference between global path planning and local path planning [87]. To adapt to the farmland operational environment, current optimization strategies employed by research scholars to enhance algorithms typically include path-planning algorithms that minimize time and distance costs; algorithms that accommodate various shapes, sizes, and locations of obstacles; and algorithms that impose constraints on the speed and energy consumption of agricultural machinery. By summarizing the published literature, we find that global path-planning algorithms can be used for local path planning after improvement, and local path-planning algorithms can be used for global path planning after optimization. Therefore, focusing on integrating global path-planning algorithms with local path-planning algorithms can enhance the realization of complete-coverage path planning.

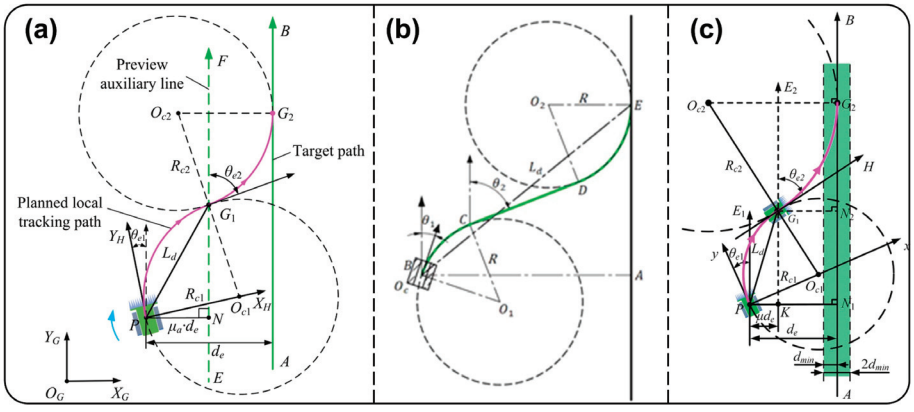


Figure 5. A localized path-planning method based on the arc tangent method. (**a**) Schematic representation of the improved pure pursuit model using the two-segment tangent method. (**b**) Schematic diagram of three-tangent localized path planning. (**c**) Schematic of adaptive localized path planning.

3.2. Bionics-Based Path-Planning Algorithms

The main bionics-based path-planning algorithms are the genetic algorithm (GA), the particle swarm optimization algorithm (PSO), and ant colony algorithm (ACO) [87]. A genetic algorithm-based path-planning optimization algorithm is a heuristic search optimization algorithm designed to find the optimal path by simulating natural selection and genetic mechanisms [88–90]. The path-planning optimization algorithm, based on a genetic algorithm, can effectively search the complex path space to find the best path solution that satisfies the specific optimization objective, which is especially suitable for generating better paths in field environments with complex terrain and many obstacles. The particle swarm optimization algorithm is an evolutionary computational technique initially developed as a simplified model inspired by the coordinated flocking behavior of birds in flight [91]. The process of path planning based on the particle swarm optimization algorithm can be regarded as the process of finding the optimal position of many particles in the solution space. The ant colony algorithm is designed to simulate the foraging behavior of ants in nature. When applied to path planning, the algorithm utilizes a pheromone update mechanism for path searching. Ants leave pheromones on the path, and other ants choose the path based on the pheromone concentration, thus gradually finding the optimal path. The above three algorithms are all group intelligence algorithms (pseudo-code comparisons are shown in Table 4) that solve optimization problems by simulating the behavior of organisms in nature or the behavior of groups.

Table 4. Pseudo-code comparison of bionics-based path-planning algorithms.

Step	GA	PSO	ACO
Initialization	Initialize population	Initialize particles	Initialize ants
Fitness Eval.	Evaluate fitness	Evaluate fitness	Evaluate fitness
Selection	Roulette wheel selection	N/A	Select next node based on probability
Crossover	Single-point crossover	N/A	N/A
Mutation	Swap mutation	N/A	N/A
Update Ind.	Replace individual	Update velocity and position	Update pheromone
Update Best	Find best individual	Update global best	Find global best path
Iteration Loop	Repeat for max generations	Repeat for max iterations	Repeat for max iterations
Return Result	Return best individual	Return global best	Return global best path

N/A means this item is empty.

All three of the above algorithms possess global search capabilities and advantages in parallel processing [92,93], enabling them to find optimal solutions in complex spaces. A comparison of the effectiveness of bionics-based path-planning algorithms when applied to path planning for farm machinery used in large fields is shown in Table 5. However, they also suffer from complex parameter tuning, unstable convergence speed, and decreased efficiency in high-dimensional spaces. The adaptability to specific problems and the interpretability of the results also need further improvement and research. Therefore, to suit the agricultural field operational environment, current optimization strategies employed by research scholars to enhance algorithms typically encompass path-planning algorithms that minimize time and distance costs, algorithms adaptable to various shapes, sizes, and locations of obstacles, and algorithms imposing constraints on the speed and energy consumption of agricultural machinery. In dynamic environments, various real-time planning algorithms are chosen to navigate constraints such as moving targets and adapting effectively to diverse operating environments. Zhou et al. [94] proposed a traversal path-planning method by combining the Floyd algorithm with an improved genetic algorithm. The traversal path problem when agricultural equipment traverses multiple plots is adapted. Xu et al. [95] considered the kinematic model of agricultural machinery, defined the paths that satisfy the kinematic model through parametric equations, and solved the initial paths using an analytical method. Meanwhile, an improved real-time path-planning algorithm for a parametric kinematic model of agricultural machinery is proposed based on the particle swarm optimization algorithm. Zhao et al. [96] proposed a path-planning method based on an improved bio-neurodynamic approach to enhance the path-planning capability of arbitrarily shaped differential-drive agricultural robots. Through several experiments in narrow and complex environments, the method was proven to have a better path-planning ability. Zhang et al. [97] compared the hyper volume estimation algorithm, the grid-based evolutionary algorithm, the turning point-driven evolutionary algorithm, and the non-dominated sorting genetic algorithm to assess their effectiveness in path planning. Their goal was to shorten the total path of agricultural machinery operation and reduce the total turning angle. Then, they proposed a multi-objective evolutionary algorithm-based path-planning method for agricultural mobile robots. Wu et al. [98] proposed an improved adaptive ant colony optimization algorithm by introducing a heuristic mechanism for directional information, an adaptive pseudo-random transfer rule, and an unevenly distributed pheromone approach, in conjunction with the traditional ant colony algorithm. The method improves the superiority of the ant colony optimization algorithm in terms of convergence speed and global optimal solution search capability.

Table 5. Comparison of bionics-based path-planning algorithms.

Algorithm Category	Global Search Ability	Convergence Speed	Computational Complexity	Adaptability	Scalability
Genetic algorithm	★★★★	★★	★★★★	★★★★	★★★
Particle swarm optimization	★★★★	★★★★★	★★★	★★★	★★★★
Ant colony algorithm	★★★★★	★★★	★★★★	★★★★	★★★★

★ represents the degree.

Currently, significant research progress has been made in applying complete-coverage path-planning techniques in unstructured farmland environments. This research primarily focuses on traditional algorithms (such as the A * algorithm, Dijkstra's algorithm, geometric methods, dynamic window approaches, etc.) and population intelligence algorithms (including the ant colony optimization algorithm, the genetic algorithm, the particle swarm optimization algorithm, etc.). By addressing global path planning and local path-planning issues, operational efficiency, farmland coverage, and operational safety have been enhanced.

4. Autonomous Operation Control Technology

Autonomous operation control technology is mainly based on the upper computer-specified commands to control the movement and operation of intelligent agricultural equipment. Taking the combined harvester shown in Figure 6 as an example, the autonomous operation control system consists of a navigation sensor unit, a user interface unit, a height sensor for the cutting table, an angle sensor, a tracking control unit, and a hydraulic actuator unit. Currently, common research in motion control primarily focuses on navigation tracking control. Research in operation control primarily focuses on equipment controlling for tasks such as ploughing, planting, management, and harvesting operations. Autonomous operation control is committed to improving the ability of self-driving farm equipment to cope with complex farmland environments and to solving the uncertainty problem brought abought by the low feedback and large inertia of farm machinery in operation control.

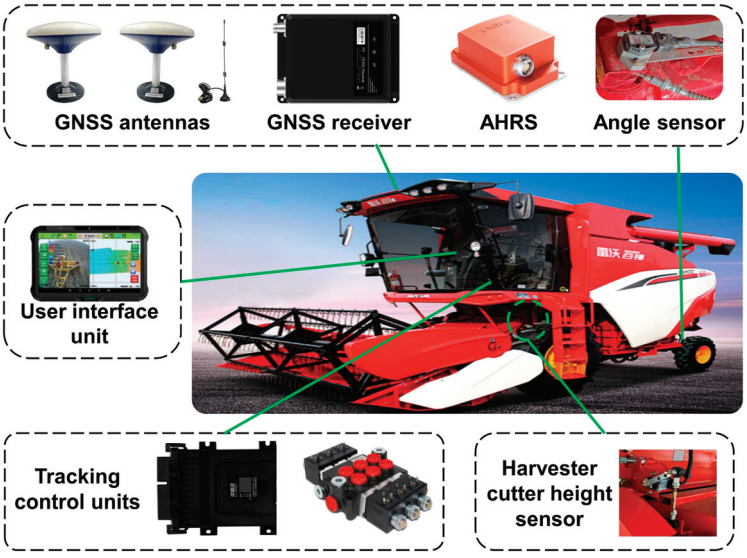


Figure 6. Autonomous operation control system and components for combine harvesters.

Motion control is the basis of autonomous operation control, and domestic and international research on the motion control of intelligent agricultural machines mainly includes PID control methods, fuzzy control methods, model predictive control methods, optimal control methods, and so on. For instance, to address the nonlinear characteristics of the tractor's longitudinal driveline system, Wang et al. [99] linearized the tractor's longitudinal dynamics model using an inverse modelling approach and designed a sliding-mode variable-structure controller to mitigate the impact of external disturbances on acceleration. Miao et al. [100] decoupled the control system into longitudinal and transverse motion control and designed a velocity controller based on the PID control algorithm. Wang et al. [101] designed a large-angle steering control algorithm based on the instantaneous center of rotation of the tracked vehicle for the steering characteristics of tracked agricultural machines based on unilateral braking. It enables the tracked vehicle to steer directly to the target heading. Navigation path tracking is the core technology to realize automatic navigation, and it is also the current research hotspot in motion control. In recent years, research scholars have applied model-free intelligent control methods (e.g., reinforcement learning, particle swarm optimization, neural networks, etc.) to navigation path tracking and obtained better results. Shan et al. [102] designed a reinforcement learning model to integrate the PID controller with the PP controller to handle the tracking error and obtain better path

tracking accuracy. Zhang et al. [103] proposed a two-depth Q-network-based vehicle path tracking control method for mobile robots to achieve the accurate tracking of straight paths and smooth transitions between polygonal trajectories. This method considers driving speed and steering rate constraints to better approximate real-world scenarios.

Operation control is a direct part of ensuring the quality of an unmanned operation of field agricultural equipment. Domestic and international research on unmanned operation control of field agricultural equipment primarily focuses on equipment controlling for ploughing, harvesting, sowing, and levelling operations. Wang et al. [104] developed a tillage speed and slip ratio switching control system for wheeled electric tractors, which can effectively control the speed and slip ratio under different tillage resistance, and has better results in terms of tillage efficiency and tillage depth stability. To solve the problem of a low level of automated control of the cutting platform height, Tan et al. [105] designed a harvester cutting platform height adaptive system based on PID control technology and tested the impact of activating the adaptive system at different forward speeds and operating modes on the operating effect. The results show that the system meets the harvester's operational needs by adjusting the height of the cutting platform. Xue et al. [106] designed a high-precision sowing depth control system for wheat planters based on fuzzy PID control technology using real-time sowing depth as the feedback input, which ensured the stability and consistency of sowing depth. Jing et al. [107] designed an adaptive PID navigation control method considering the side-slip estimation to enhance the operational stability of the unmanned farm grader, considering the side-slip phenomenon that occurs during operation. At present, the research on motion control and operation control for intelligent agricultural machinery and equipment has yielded significant results.

5. Conclusions and Prospection

5.1. Conclusions

Agricultural equipment achieves the perception of the complex and changing operating environment around it by carrying onboard environmental sensing sensors. It carries out autonomous path planning and navigation according to the preset operation tasks and map information and carries out the operation control of the operation equipment, thus realizing the unmanned operation of large-scale agricultural equipment. To improve the level of environmental sensing, path planning, and autonomous operation control of large-scale agricultural equipment, researchers have carried out many studies. This review systematically examines the progress of key technologies for the unmanned operation of field agricultural equipment from three aspects: onboard environmental sensing technology, coverage path-planning technology, and autonomous operation control technology. Although the unmanned operation technology of field agricultural equipment has a wide range of application prospects, it still faces a series of challenges and problems in practical application. The challenges of robustness of environment perception, practicality of path-planning methods, and efficiency of autonomous operation control methods need to be further improved. Through the systematic review, the key conclusions are drawn as follows:

(1) Single-sensor onboard environmental sensing methods are no longer adequate to meet the demands of unmanned operations. Traditional vision methods are affected by environmental factors such as light and weather, resulting in insufficient stability and accuracy under complex farmland conditions. Although LiDAR can provide high precision distance information, it is susceptible to errors when interfered with by particulate matter in agricultural fields. Therefore, the advantages of multiple sensors, such as vision, LiDAR, and millimeter-wave radar, should be combined. Their characteristics in information acquisition and environmental adaptability can be comprehensively utilized to enhance the accuracy and robustness of farmland boundary detection, navigation line extraction, and obstacle detection.

(2) Classical complete-coverage path-planning algorithms such as A * and Dijkstra are suitable for planning in static environments but can result in unsmooth and inefficient

paths in complex farmlands. The introduction of methods based on artificial potential fields, sampling, and curve fitting can improve the accuracy and adaptability of path planning. In addition, intelligent bionics-based algorithms, such as the genetic algorithm, particle swarm optimization, and the ant colony algorithm, which optimize path selection by simulating biological behavior, have certain advantages in finding the globally optimal path in complex environments. Although complete-coverage path-planning techniques have been well developed, the related algorithms are deficient in accuracy and timeliness when dealing with terrain changes and the real-time requirements of dynamic farmland environments.

(3) Research on motion control focuses on speed, steering, and navigation tracking, using PID, fuzzy control, model prediction, optimal control, and other methods. Operational control encompasses scenarios such as ploughing, harvesting, seeding, and grading operations, primarily enhancing operational quality and efficiency through PID adaptive systems and fuzzy control techniques. While these research methods have significantly enhanced the unmanned operation capabilities of large-scale agricultural equipment, they still inadequately address the low feedback and high inertia issues encountered by medium- to large-size agricultural machinery during field operations. Autonomous operation systems exhibit poor robustness under varying ground conditions, significant disparities in tracking control performance, and limited adaptive capabilities to different ground conditions.

5.2. Prospection

Large-scale fields are the primary production zones for food crops, and the research on agricultural equipment for these vast terrains is a critical issue that demands urgent attention in advancing large-scale agriculture. The autonomous operation of agricultural machinery in large farmland areas is a pivotal strategy to enhance production efficiency, reduce labor intensity, and mitigate the impact of an aging population on agriculture. A comprehensive overview of the research progress on key technologies for the autonomous operation of large-scale farmland machinery will facilitate the systematic advancement of agricultural equipment in these vast areas. Regarding onboard environmental sensing technology, multi-sensor fusion sensing technology has demonstrated significant advantages in environmental sensing. However, existing research still faces the following shortcomings: the complexity of data fusion algorithms results in inadequate real-time performance, interference, and signal inconsistency among sensors. These issues negatively impact perception accuracy, and there is poor adaptability to harsh environments. Therefore, future research can focus on the following two aspects: first, optimizing data fusion strategies and algorithms to enhance real-time and processing efficiency, to ensure rapid response in complex farmland environments; second, improving sensor anti-interference capability and signal consistency, adopting advanced calibration and compensation techniques to enhance sensing accuracy and adaptability to harsh environments.

In terms of complete-coverage path-planning technology, current research mainly focuses on single-field complete-coverage route planning technology. Existing research has improved the real-time and adaptive nature of path planning by improving the relevant algorithms, thus increasing operational efficiency, farmland coverage, and operational safety. However, there have been relatively few studies on complete-coverage path planning for multiple fields. To further enhance the efficiency of unmanned operations for large-scale agricultural equipment, future research should prioritize the following aspects: First, develop efficient path-planning algorithms applicable to multi-field environments to ensure a smooth transition and path optimization between different farmlands. Second, with the gradual realization of intelligent operation of agricultural machinery, the limitations of single-machine agricultural operations are gradually becoming evident. Research should focus on developing multi-machine intelligent scheduling and task allocation systems tailored to diverse operational requirements across multiple fields, optimizing operation sequencing and resource allocation. Leveraging the collective intelligence of groups of agricultural machines can significantly enhance overall agricultural production efficiency.

In terms of autonomous operational control technology, navigation accuracy and operational quality have become critical factors in the effectiveness of operational systems. Therefore, further research should investigate the variations in motion trajectories of field agricultural equipment under different ground conditions and implement targeted compensatory control measures to enhance the robustness, adaptability, and operational accuracy of autonomous operation control systems.

Author Contributions: Conceptualization, W.W., M.X. and H.W.; investigation, W.D., H.W. and C.Z.; writing—original draft preparation, W.W. and W.D.; writing—review and editing, W.W., C.Z. and Y.Z.; funding acquisition, M.X. and G.G. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by The National Key Research and Development Program of China, grant number 2022YFD2001604.

Data Availability Statement: No data were used for the research described in this article.

Conflicts of Interest: Authors Hui Wang and Cheng Zhai were employed by the company Weichai Lovol Intelligent Agricultural Technology Co., Ltd. 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.

References

1. Charania, I.; Li, X. Smart farming: Agriculture's shift from a labor intensive to technology native industry. *Internet Things* **2020**, *9*, 100142. [CrossRef]
2. Ren, C.; Zhou, X.; Wang, C.; Guo, Y.; Diao, Y.; Shen, S.; Reis, S.; Li, W.; Xu, J.; Gu, B. Ageing threatens sustainability of smallholder farming in China. *Nature* **2023**, *616*, 96–103. [CrossRef]
3. Liu, J.; Fang, Y.; Wang, G.; Liu, B.; Wang, R. The aging of farmers and its challenges for labor-intensive agriculture in China: A perspective on farmland transfer plans for farmers' retirement. *J. Rural. Stud.* **2023**, *100*, 103013. [CrossRef]
4. Chen, R.; Xu, P.; Song, P.; Wang, M.; He, J. China has faster pace than Japan in population aging in next 25 years. *Biosci. Trends* **2019**, *13*, 287–291. [CrossRef] [PubMed]
5. Yu, Y.; Zhang, J.; Zhang, K.; Xu, D.; Qi, Y.; Deng, X. The impacts of farmer ageing on farmland ecological restoration technology adoption: Empirical evidence from rural China. *J. Clean. Prod.* **2023**, *430*, 139648. [CrossRef]
6. Cui, Z.; Zhang, H.; Chen, X.; Zhang, C.; Ma, W.; Huang, C.; Zhang, W.; Mi, G.; Miao, Y.; Li, X.; et al. Pursuing sustainable productivity with millions of smallholder farmers. *Nature* **2018**, *555*, 363–366. [CrossRef] [PubMed]
7. Zhang, Q.; Chu, Y.; Xue, Y.; Ying, H.; Chen, X.; Zhao, Y.; Ma, W.; Ma, L.; Zhang, J.; Yin, Y.; et al. Outlook of China's agriculture transforming from smallholder operation to sustainable production. *Glob. Food Secur.* **2020**, *26*, 100444. [CrossRef]
8. Yao, Z.; Zhao, C.; Zhang, T. Agricultural machinery automatic navigation technology. *Iscience* **2024**, *27*, 108714. [CrossRef] [PubMed]
9. Shuai, L.; Li, Z.; Chen, Z.; Luo, D.; Mu, J. A research review on deep learning combined with hyperspectral Imaging in multiscale agricultural sensing. *Comput. Electron. Agric.* **2024**, *217*, 108577. [CrossRef]
10. Ariza-Sentís, M.; Vélez, S.; Martínez-Peña, R.; Baja, H.; Valente, J. Object detection and tracking in Precision Farming: A systematic review. *Comput. Electron. Agric.* **2024**, *219*, 108757. [CrossRef]
11. Zhang, S.; Liu, Y.; Xiong, K.; Tian, Y.; Du, Y.; Zhu, Z.; Du, M.; Zhai, Z. A review of vision-based crop row detection method: Focusing on field ground autonomous navigation operations. *Comput. Electron. Agric.* **2024**, *222*, 109086. [CrossRef]
12. Liu, H.; Li, K.; Ma, L.; Meng, Z. Headland Identification and Ranging Method for Autonomous Agricultural Machines. *Agriculture* **2024**, *14*, 243. [CrossRef]
13. Liang, C.; Pan, K.; Zhao, M.; Lu, M. Multi-Node Path Planning of Electric Tractor Based on Improved Whale Optimization Algorithm and Ant Colony Algorithm. *Agriculture* **2023**, *13*, 586. [CrossRef]
14. Wan, H.; Ou, Y.; Guan, X.; Jiang, R.; Zhou, Z.; Luo, X. Review of the perception technologies for unmanned agricultural machinery operating environment. *Trans. CSAE* **2024**, *40*, 1–18. [CrossRef]
15. Bai, Y.; Zhang, B.; Xu, N.; Zhou, J.; Shi, J.; Diao, Z. Vision-based navigation and guidance for agricultural autonomous vehicles and robots: A review. *Comput. Electron. Agric.* **2023**, *205*, 107584. [CrossRef]
16. Barbedo, J.G.A. Data Fusion in Agriculture: Resolving Ambiguities and Closing Data Gaps. *Sensors* **2022**, *22*, 2285. [CrossRef]
17. Höffmann, M.; Patel, S.; Büskens, C. Optimal guidance track generation for precision agriculture: A review of coverage path planning techniques. *J. Field Robot.* **2024**, *41*, 823–844. [CrossRef]
18. Sánchez-Ibáñez, J.R.; Pérez-del-Pulgar, C.J.; García-Cerezo, A. Path Planning for Autonomous Mobile Robots: A Review. *Sensors* **2021**, *21*, 7898. [CrossRef]
19. Zhang, G.; Ji, C.; Wu, Q.; Liu, H.; Zhou, Y.; Fu, J. Study on path planning of mechanized harvesting of ratoon rice in the first season based on the capacitated arc routing problem model. *Front. Plant Sci.* **2022**, *13*, 963307. [CrossRef]

20. Guan, Z.; Li, Y.; Mu, S.; Zhang, M.; Jiang, T.; Li, H.; Wang, G.; Wu, C. Tracing algorithm and control strategy for crawler rice combine harvester auxiliary navigation system. *Biosyst. Eng.* **2021**, *211*, 50–62. [CrossRef]
21. An, G.; Zhong, Z.; Yang, S.; Yang, L.; Jin, C.; Du, J.; Yin, X. EASS: An automatic steering system for agricultural wheeled vehicles using fuzzy control. *Comput. Electron. Agric.* **2024**, *217*, 108544. [CrossRef]
22. Zhao, X.; Lu, E.; Tang, Z.; Luo, C.; Xu, L.; Wang, H. Trajectory prediction method for agricultural tracked robots based on slip parameter estimation. *Comput. Electron. Agric.* **2024**, *222*, 109057. [CrossRef]
23. Wang, Q.; Meng, Z.; Wen, C.; Qin, W.; Wang, F.; Zhang, A.; Zhao, C.; Yin, Y. Grain combine harvester header profiling control system development and testing. *Comput. Electron. Agric.* **2024**, *223*, 109082. [CrossRef]
24. Cui, Z.; Hu, J.; Yu, Y.; Cao, G.; Zhang, H.; Chai, X.; Chen, H.; Xu, L. Automatic grain unloading method for track-driven rice combine harvesters based on stereo vision. *Comput. Electron. Agric.* **2024**, *220*, 108917. [CrossRef]
25. Rovira-Mas, F.; Saiz-Rubio, V.; Cuenca-Cuenca, A. Augmented Perception for Agricultural Robots Navigation. *IEEE Sens. J.* **2021**, *21*, 11712–11727. [CrossRef]
26. Xie, D.; Chen, L.; Liu, L.; Chen, L.; Wang, H. Actuators and Sensors for Application in Agricultural Robots: A Review. *Machines* **2022**, *10*, 913. [CrossRef]
27. Lu, H.; Wang, H.; Ma, Z.; Ren, Y.; Fu, W.; Shan, Y.; Hu, S.; Zhang, G.; Meng, Z. Farmland boundary extraction based on the AttMobile-DeeplabV3+ network and least squares fitting of straight lines. *Front. Plant Sci.* **2023**, *14*, 1228590. [CrossRef]
28. Hong, Q.; Zhu, Y.; Liu, W.; Ren, T.; Shi, C.; Lu, Z.; Yang, Y.; Deng, R.; Qian, J.; Tan, C. A segmentation network for farmland ridge based on encoder-decoder architecture in combined with strip pooling module and ASPP. *Front. Plant Sci.* **2024**, *15*, 1328075. [CrossRef]
29. Yang, Y.; Shen, X.; An, D.; Han, H.; Tang, W.; Wang, Y.; Yang, Y.; Ma, Q.; Chen, L. Crop Row Detection Algorithm Based on 3-D LiDAR: Suitable for Crop Row Detection in Different Periods. *IEEE Trans. Instrum. Meas.* **2024**, *73*, 8503413. [CrossRef]
30. Campos, Y.; Sossa, H.; Pajares, G. Spatio-temporal analysis for obstacle detection in agricultural videos. *Appl. Soft Comput.* **2016**, *45*, 86–97. [CrossRef]
31. Pini, M.; Marucco, G.; Falco, G.; Nicola, M.; De Wilde, W. Experimental Testbed and Methodology for the Assessment of RTK GNSS Receivers Used in Precision Agriculture. *IEEE Access* **2020**, *8*, 14690–14703. [CrossRef]
32. Jiménez-Martínez, M.J.; Farjas-Abadia, M.; Quesada-Olmo, N. An Approach to Improving GNSS Positioning Accuracy Using Several GNSS Devices. *Remote Sens.* **2021**, *13*, 1149. [CrossRef]
33. Perez-Ruiz, M.; Martínez-Guanter, J.; Upadhyaya, S.K. Chapter 15—High-precision GNSS for agricultural operations. In *GPS and GNSS Technology in Geosciences*; Elsevier: Amsterdam, The Netherlands, 2021; pp. 299–335. [CrossRef]
34. Chen, Y.; Li, G.; Zhang, X.; Jia, J.; Zhou, K.; Wu, C. Identifying field and road modes of agricultural Machinery based on GNSS Recordings: A graph convolutional neural network approach. *Comput. Electron. Agric.* **2022**, *198*, 107082. [CrossRef]
35. Moeller, R.; Deemyad, T.; Sebastian, A. Autonomous Navigation of an Agricultural Robot Using RTK GPS and Pixhawk. In Proceedings of the 2020 Intermountain Engineering, Technology and Computing (IETC), Orem, UT, USA, 2–3 October 2020; IEEE: Piscataway, NJ, USA, 2020; pp. 1–6. [CrossRef]
36. Ren, H.; Wu, J.; Lin, T.; Yao, Y.; Liu, C. Research on an Intelligent Agricultural Machinery Unmanned Driving System. *Agriculture* **2023**, *13*, 1907. [CrossRef]
37. Huang, Y.; Fu, J.; Xu, S.; Han, T.; Liu, Y. Research on Integrated Navigation System of Agricultural Machinery Based on RTK-BDS/INS. *Agriculture* **2022**, *12*, 1169. [CrossRef]
38. LeVoir, S.J.; Farley, P.A.; Sun, T.; Xu, C. High-Accuracy Adaptive Low-Cost Location Sensing Subsystems for Autonomous Rover in Precision Agriculture. *IEEE Open J. Ind. Appl.* **2020**, *1*, 74–94. [CrossRef]
39. Liang, X.; Huang, Z.; Qin, H. Differential positioning based on the orthogonal transformation algorithm with GNSS multi-system. *GPS Solut.* **2018**, *22*, 89. [CrossRef]
40. Yue, B.; Zhang, Z.; Zhang, W.; Luo, X.; Zhang, G.; Huang, H.; Wu, X.; Bao, K.; Peng, M. Design of an Automatic Navigation and Operation System for a Crawler-Based Orchard Sprayer Using GNSS Positioning. *Agronomy* **2024**, *14*, 271. [CrossRef]
41. Kaivosoja, J.; Linkolehto, R. GNSS error simulator for farm machinery navigation development. *Comput. Electron. Agric.* **2015**, *119*, 166–177. [CrossRef]
42. Lee, K.; Choi, H.; Kim, J. Development of Path Generation and Algorithm for Autonomous Combine Harvester Using Dual GPS Antenna. *Sensors* **2023**, *23*, 4944. [CrossRef]
43. Vulpi, F.; Marani, R.; Petitti, A.; Reina, G.; Milella, A. An RGB-D multi-view perspective for autonomous agricultural robots. *Comput. Electron. Agric.* **2022**, *202*, 107419. [CrossRef]
44. Galati, R.; Mantriota, G.; Reina, G. RoboNav: An Affordable Yet Highly Accurate Navigation System for Autonomous Agricultural Robots. *Robotics* **2022**, *11*, 99. [CrossRef]
45. Valente, D.S.M.; Momin, A.; Grift, T.; Hansen, A. Accuracy and precision evaluation of two low-cost RTK global navigation satellite systems. *Comput. Electron. Agric.* **2020**, *168*, 105142. [CrossRef]
46. Jing, Y.; Li, Q.; Ye, W.; Liu, G. Development of a GNSS/INS-based automatic navigation land levelling system. *Comput. Electron. Agric.* **2023**, *213*, 108187. [CrossRef]
47. Li, S.; Zhang, M.; Ji, Y.; Zhang, Z.; Cao, R.; Chen, B.; Li, H.; Yin, Y. Agricultural machinery GNSS/IMU-integrated navigation based on fuzzy adaptive finite impulse response Kalman filtering algorithm. *Comput. Electron. Agric.* **2021**, *191*, 106524. [CrossRef]

48. Hong, Z.; Li, Y.; Lin, H.; Gong, L.; Liu, C. Field Boundary Distance Detection Method in Early Stage of Planting Based on Binocular Vision. *Trans. CSAM* **2022**, *53*, 27–33. [CrossRef]
49. Wang, Q.; Liu, H.; Yang, P.; Meng, Z. Detection Method of Headland Boundary Line Based on Machine Vision. *Trans. CSAM* **2020**, *51*, 18–27. [CrossRef]
50. Li, D.; Li, B.; Long, S.; Feng, H.; Wang, Y.; Wang, J. Robust detection of headland boundary in paddy fields from continuous RGB-D images using hybrid deep neural networks. *Comput. Electron. Agric.* **2023**, *207*, 107713. [CrossRef]
51. Yang, L.; Xu, Y.; Li, Y.; Chang, M.; Chen, Z.; Lan, Y.; Wu, C. Real-Time field road freespace extraction for agricultural machinery autonomous driving based on LiDAR. *Comput. Electron. Agric.* **2023**, *211*, 108028. [CrossRef]
52. Turner, R.; Panciera, R.; Tanase, M.A.; Lowell, K.; Hacker, J.M.; Walker, J.P. Estimation of soil surface roughness of agricultural soils using airborne LiDAR. *Remote Sens. Environ.* **2014**, *140*, 107–117. [CrossRef]
53. Wang, X.; Shu, L.; Han, R.; Yang, F.; Gordon, T.; Wang, X.; Xu, H. A Survey of Farmland Boundary Extraction Technology Based on Remote Sensing Images. *Electronics* **2023**, *12*, 1156. [CrossRef]
54. Waldner, F.; Diakogiannis, F.I. Deep learning on edge: Extracting field boundaries from satellite images with a convolutional neural network. *Remote Sens. Environ.* **2020**, *245*, 111741. [CrossRef]
55. Hou, W.; Zhou, C.; Cheng, Y.; Wang, Y.; Liu, L.; Qin, K. Path recognition method of orchard ridges based on lightweight U-Net. *Trans. CSAM* **2024**, *55*, 16–27. [CrossRef]
56. Gong, J.; Wang, X.; Zhang, Y.; Lan, Y.; Mostafa, K. Navigation line extraction based on root and stalk composite locating points. *Comput. Electr. Eng.* **2021**, *92*, 107115. [CrossRef]
57. Diao, Z.; Guo, P.; Zhang, B.; Zhang, D.; Yan, J.; He, Z.; Zhao, S.; Zhao, C.; Zhang, J. Navigation line extraction algorithm for corn spraying robot based on improved YOLOv8s network. *Comput. Electron. Agric.* **2023**, *212*, 108049. [CrossRef]
58. Yang, Y.; Mang, Q.; Chen, Z.; Wen, X.; Zhang, G.; Zhang, T.; Dong, X.; Chen, L. Real-time extraction of the navigation lines between sugarcane ridges using LiDAR. *Trans. CSAE* **2022**, *38*, 178–185. [CrossRef]
59. Zhou, H.; Yang, Y.; Liu, Y.; Ma, R.; Zhang, F.; Zang, Q. Real-time Extraction of Navigation Line Based on LiDAR. *Trans. CSAM* **2023**, *54*, 9–17. [CrossRef]
60. Wang, S.; Song, J.; Qi, P.; Yuan, C.; Wu, H.; Zhang, L.; Liu, W.; Liu, Y.; He, X. Design and development of orchard autonomous navigation spray system. *Front. Plant Sci.* **2022**, *13*, 960686. [CrossRef] [PubMed]
61. Liu, X.; Zhang, C.; Zhang, H.; Yang, S.; Jiang, S.; Zheng, Y.; Su, D.; Wan, C. Inter-row automatic navigation method by combining least square and SVM in forestry. *Trans. CSAE* **2021**, *37*, 157–164. [CrossRef]
62. He, Y.; Jiang, H.; Fang, H.; Wang, Y.; Liu, Y. Research progress of intelligent obstacle detection methods of vehicles and their application on agriculture. *Trans. Chin. Soc. Agric. Eng.* **2018**, *34*, 21–32. [CrossRef]
63. Li, Y.; Li, M.; Qi, J.; Zhou, D.; Zou, Z.; Liu, K. Detection of typical obstacles in orchards based on deep convolutional neural network. *Comput. Electron. Agric.* **2021**, *181*, 105932. [CrossRef]
64. Xue, J.; Cheng, F.; Li, Y.; Song, Y.; Mao, T. Detection of Farmland Obstacles Based on an Improved YOLOv5s Algorithm by Using CIoU and Anchor Box Scale Clustering. *Sensors* **2022**, *22*, 1790. [CrossRef] [PubMed]
65. Xu, H.; Li, S.; Ji, Y.; Cao, R.; Zhang, M. Dynamic obstacle detection based on panoramic vision in the moving state of agricultural machineries. *Comput. Electron. Agric.* **2021**, *184*, 106104. [CrossRef]
66. Du, X.; Li, Z.; Ma, Z.; Yang, Z.; Wang, D. UAV Field Obstacle Detection Based on Spatial Attention and Deformable Convolution. *Trans. CSAM* **2023**, *54*, 275–283. [CrossRef]
67. Shang, Y.; Zhang, G.; Meng, Z.; Wang, H.; Su, C.; Song, Z. Field Obstacle Detection Method of 3D LiDAR Point Cloud Based on Euclidean Clustering. *Trans. CSAM* **2022**, *53*, 23–32. [CrossRef]
68. Wang, T.; Wang, N.; Xiao, J.; Miao, Y.; Sun, Y.; Li, H.; Zhang, M. One-shot domain adaptive real-time 3D obstacle detection in farmland based on semantic-geometry-intensity fusion strategy. *Comput. Electron. Agric.* **2023**, *214*, 108264. [CrossRef]
69. Xue, J.; Cheng, F.; Wang, B.; Li, Y.; Ma, Z.; Chu, Y. Method for Millimeter Wave Radar Farm Obstacle Detection Based on Invalid Target Filtering. *Trans. CSAM* **2023**, *54*, 233–240. [CrossRef]
70. Ball, D.; Upcroft, B.; Wyeth, G.; Corke, P.; English, A.; Ross, P.; Patten, T.; Fitch, R.; Sukkarieh, S.; Bate, A. Vision-based Obstacle Detection and Navigation for an Agricultural Robot. *J. Field Robot.* **2016**, *33*, 1107–1130. [CrossRef]
71. Lv, P.; Wang, B.; Cheng, F.; Xue, J. Multi-Objective Association Detection of Farmland Obstacles Based on Information Fusion of Millimeter Wave Radar and Camera. *Sensors* **2022**, *23*, 230. [CrossRef]
72. Cai, D. Research on Autonomous Operation Perception Technology in Unstructured Farmland Environment. Master's Thesis, Shanghai Jiao Tong University, Shanghai, China, 2020.
73. Kragh, M.; Underwood, J. Multimodal obstacle detection in unstructured environments with conditional random fields. *J. Field Robot.* **2020**, *37*, 53–72. [CrossRef]
74. Liu, L.; Wang, X.; Liu, H.; Li, J.; Wang, P.; Yang, X. A Full-Coverage Path Planning Method for an Orchard Mower Based on the Dung Beetle Optimization Algorithm. *Agriculture* **2024**, *14*, 865. [CrossRef]
75. Shen, Y.; Liu, Z.; Liu, H.; Du, W. Orchard Spray Robot Planning Algorithm Based on Multiple Constraints. *Trans. CSAM* **2023**, *54*, 56–67. [CrossRef]
76. Xu, X.; Zeng, J.; Zhao, Y.; Lü, X. Research on global path planning algorithm for mobile robots based on improved A*. *Expert Syst. Appl.* **2024**, *243*, 122922. [CrossRef]

77. Feng, Z.; Zhou, L.; Qi, J.; Hong, S. DBVS-APF-RRT*: A global path planning algorithm with ultra-high speed generation of initial paths and high optimal path quality. *Expert Syst. Appl.* **2024**, *249*, 123571. [CrossRef]
78. Kong, F.; Liu, B.; Han, X.; Yi, L.; Sun, H.; Liu, J.; Liu, L.; Lan, Y. Path Planning Algorithm of Orchard Fertilization Robot Based on Multi-Constrained Bessel Curve. *Agriculture* **2024**, *14*, 979. [CrossRef]
79. Wu, Z.; Dai, J.; Jiang, B.; Karimi, H.R. Robot Path Planning Based on Artificial Potential Field with Deterministic Annealing. *ISA Trans.* **2023**, *138*, 74–87. [CrossRef] [PubMed]
80. Cui, J.; Zhang, X.; Fan, X.; Feng, W.; Li, P.; Wu, Y. Path planning of autonomous agricultural machineries in complex rural road. *J. Eng.* **2020**, *6*, 239–245. [CrossRef]
81. Boryga, M.; Kołodziej, P.; Gołacki, K. Application of Polynomial Transition Curves for Trajectory Planning on the Headlands. *Agriculture* **2020**, *10*, 144. [CrossRef]
82. He, Y.; Zhou, J.; Sun, J.; Jia, H.; Gemechu, T.-T. Traveling control method adapted to different paddy ground conditions with feedforward compensation for crawler combine harvester based on online tracking error prediction. *Comput. Electron. Agric.* **2024**, *220*, 108853. [CrossRef]
83. He, Y.; Zhou, J.; Yuan, L.; Zheng, P.; Liang, Z. Local Tracking Path Planning Based on Steering Characteristics of Crawler-type Combine Harvester. *Trans. CSAM* **2022**, *53*, 13–21. [CrossRef]
84. He, Y.; Zhou, J.; Sun, J.; Jia, H.; Liang, Z.; Awuah, E. An adaptive control system for path tracking of crawler combine harvester based on paddy ground conditions identification. *Comput. Electron. Agric.* **2023**, *210*, 107948. [CrossRef]
85. Qi, Z.; Zhou, J.; Ye, Z.; Zhou, X.; Yang, H. Local Path Planning Method in Automatic Walking of Single-side Brake Steering Track Chassis. *Trans. CSAM* **2024**, *55*, 52–59. [CrossRef]
86. Yang, Y.; Li, Y.; Wen, X.; Zhang, G.; Ma, Q.; Cheng, S.; Qi, J.; Xu, L.; Chen, L. An optimal goal point determination algorithm for automatic navigation of agricultural machinery: Improving the tracking accuracy of the Pure Pursuit algorithm. *Comput. Electron. Agric.* **2022**, *194*, 106760. [CrossRef]
87. Liu, L.; Wang, X.; Yang, X.; Liu, H.; Li, J.; Wang, P. Path planning techniques for mobile robots: Review and prospect. *Expert Syst. Appl.* **2023**, *227*, 120254. [CrossRef]
88. Kala, R. Multi-robot path planning using co-evolutionary genetic programming. *Expert Syst. Appl.* **2012**, *39*, 3817–3831. [CrossRef]
89. Ab Wahab, M.N.; Nazir, A.; Khalil, A.; Ho, W.J.; Akbar, M.F.; Noor, M.H.M.; Mohamed, A.S.A. Improved genetic algorithm for mobile robot path planning in static environments. *Expert Syst. Appl.* **2024**, *249*, 123762. [CrossRef]
90. Shang, Y.; Shen, J.; Wei, W.; Zheng, B. Optimization of Ball Mill Cylinder Structure Based on Response Surface Optimization Module and Multi-objective Genetic Algorithm. *J. Mech. Sci. Technol.* **2024**, *38*, 3631–3640. [CrossRef]
91. Hilli, A.A.; Al-Ibadi, M.; Alfadhel, A.M.; Abdulshaheed, S.H.; Hadi, A.H. Optimal path finding in stochastic quasi-dynamic environments using particle swarm optimization. *Expert Syst. Appl.* **2021**, *186*, 115706. [CrossRef]
92. Linker, R.; Blass, T. Path-planning algorithm for vehicles operating in orchards. *Biosyst. Eng.* **2008**, *101*, 152–160. [CrossRef]
93. Reda, M.; Onsy, A.; Haikal, A.Y.; Ghanbari, A. Path planning algorithms in the autonomous driving system: A comprehensive review. *Robot. Auton. Syst.* **2024**, *174*, 104630. [CrossRef]
94. Zhou, L.; Liu, T.; Lu, J. Traversal Path Planning for Farmland in Hilly Areas based on Floyd and Im-proved Genetic Algorithm. *Smart Agric.* **2023**, *4*, 45–57. [CrossRef]
95. Xu, L.; You, J.; Yuan, H. Real-Time Parametric Path Planning Algorithm for Agricultural Machinery Kinematics Model Based on Particle Swarm Optimization. *Agriculture* **2023**, *13*, 1960. [CrossRef]
96. Zhao, Z.; Liu, S.; Wei, J.; Qin, F. Improved biological neural network approach for path planning of differential drive agricultural robots with arbitrary shape. *Comput. Electron. Agric.* **2024**, *216*, 108525. [CrossRef]
97. Zhang, X.; Guo, Y.; Yang, J.; Li, D.; Wang, Y.; Zhao, R. Many-objective evolutionary algorithm based agricultural mobile robot route planning. *Comput. Electron. Agric.* **2022**, *200*, 107274. [CrossRef]
98. Wu, L.; Huang, X.; Cui, J.; Liu, C.; Xiao, W. Modified adaptive ant colony optimization algorithm and its application for solving path planning of mobile robot. *Expert Syst. Appl.* **2023**, *215*, 119410. [CrossRef]
99. Wang, Z.; Liu, Z.; Bai, X.; Gao, L. Longitudinal Acceleration Tracking Control of Tractor Cruise System. *Trans. CSAM* **2018**, *49*, 21–28. [CrossRef]
100. Miao, H.; Diao, P.; Xu, G.; Yao, W.; Song, Z.; Wang, W. Research on decoupling control for the longitudinal and lateral dynamics of a tractor considering steering delay. *Sci. Rep.* **2022**, *12*, 13997. [CrossRef]
101. Wang, F.; Yang, Q.; Zhang, Z.; Li, A.; Xu, H. Path Tracking and Turning Control Algorithm of Tracked Vehicle Based on ICR. *Trans. CSAM* **2024**, *55*, 386–395. [CrossRef]
102. Shan, Y.; Zheng, B.; Chen, L.; Chen, L.S.; Chen, D. A Reinforcement Learning-Based Adaptive Path Tracking Approach for Autonomous Driving. *IEEE T Veh. Technol.* **2020**, *69*, 10581–10595. [CrossRef]
103. Zhang, W.; Gai, J.; Zhang, Z.; Tang, L.; Liao, Q.; Ding, Y. Double-DQN based path smoothing and tracking control method for robotic vehicle navigation. *Comput. Electron. Agric.* **2019**, *166*, 104985. [CrossRef]
104. Wang, Q.; Wang, X.; Wang, W.; Song, Y.; Cui, Y. Joint control method based on speed and slip rate switching in plowing operation of wheeled electric tractor equipped with sliding battery pack. *Comput. Electron. Agric.* **2023**, *215*, 108426. [CrossRef]
105. Tan, H.; Wang, G.; Zhou, S.; Jia, H.; Qu, M.; Xiang, M.; Gao, X.; Zhou, Z.; Li, H.; Zou, Z. Design and Experiment of Header Height Adaptive Adjustment System for Maize (*Zea mays* L.) Harvester. *Sustainability* **2023**, *15*, 14137. [CrossRef]

106. Xue, B.; Zhou, L.; Niu, K.; Zheng, Y.; Bai, S.; Gui, L. Sowing Depth Control System of Wheat Planter Based on Adaptive Fuzzy PID. *Trans. CSAM* **2023**, *54*, 93–102. [CrossRef]
107. Jing, Y.; Liu, G.; Jin, Z. Adaptive Guidance Control Method with Sideslip Estimation for Land Leveler Pulling by Tractor. *Trans. CSAM* **2020**, *51*, 26–33. [CrossRef]

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.

agriculture

MDPI

Review

The Application of Artificial Intelligence Models for Food Security: A Review

Rebecca Sarku *, Ulfia A. Clemen and Thomas Clemen

Multi-Agent Research and Simulation (MARS Group), Hamburg University of Applied Sciences, Berliner Tor 7, 20099 Hamburg, Germany; ulfia.clemen@haw-hamburg.de (U.A.C.); thomas.clemen@haw-hamburg.de (T.C.)
* Correspondence: rebecca.sarku@haw-hamburg.de

Abstract: Emerging technologies associated with Artificial Intelligence (AI) have enabled improvements in global food security situations. However, there is a limited understanding regarding the extent to which stakeholders are involved in AI modelling research for food security purposes. This study systematically reviews the existing literature to bridge the knowledge gap in AI and food security, focusing on software modelling perspectives. The study found the application of AI models to examine various indicators of food security across six continents, with most studies conducted in sub-Saharan Africa. While research organisations conducting AI modelling were predominantly based in Europe or the Americas, their study communities were in the Global South. External funders also supported AI modelling research on food security through international universities and research institutes, although some collaborations with local organisations and external partners were identified. The analysis revealed three patterns in the application of AI models for food security research: (1) the exclusive utilisation of AI models to assess food security situations, (2) stakeholder involvement in some aspects of the AI modelling process, and (3) stakeholder involvement in AI modelling for food security through an iterative process. Overall, studies on AI models for food security were primarily experimental and lacked real-life implementation of the results with stakeholders. Consequently, this study concluded that research on AI, which incorporates feedback and/or the implementation of research outcomes for stakeholders, can contribute to learning and enhance the validity of the models in addressing food security challenges.

Keywords: Artificial Intelligence; food security; machine learning; Global South; funding

Citation: Sarku, R.; Clemen, U.A.; Clemen, T. The Application of Artificial Intelligence Models for Food Security: A Review. *Agriculture* **2023**, *13*, 2037. https://doi.org/10.3390/agriculture13102037

Academic Editors: Jin Yuan and Zichen Huang

Received: 27 September 2023
Revised: 6 October 2023
Accepted: 11 October 2023
Published: 23 October 2023

Copyright: © 2023 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (https://creativecommons.org/licenses/by/4.0/).

1. Background

The pursuit of a sustainable food system has prompted the adoption of innovative technologies like Artificial Intelligence (AI) to enhance food security. Artificial Intelligence (AI) encompasses a wide range of disciplines within computer science, with the aim of constructing intelligent machines capable of executing tasks typically demanding human intelligence [1]. AI has become imperative due to the uncertainties associated with the intricate and dynamic relationship among social, economic, and environmental factors that underlie food security [2,3]. For instance, a report by the FAO, IFAD, UNICEF, WFP, and WHO [4] indicates that global food insecurity is projected to worsen due to rapid population growth, overexploitation, the depletion of natural resources, and unprecedented climate change. About 70% more food will be needed by 2050 to feed the growing population, while the global food consumption in calories from the use of agricultural commodities is projected to increase by 1.3% per year over the next decade [5]. Food for developing countries is produced by 80% of smallholder farmers who rely on simple technologies [6]. The rise in agricultural input prices experienced over the last two years, particularly the rising fertiliser costs, can lead to higher food prices, and this raises concern about global food security [7]. Stakeholders in the food and agricultural sector face the challenge of optimising their operations to minimise losses and costs while maximising yields. This

challenge encompasses factors, such as low crop yields, losses due to weather events (e.g., droughts, floods, and frost), pest and disease incidences, post-harvest losses during storage and transportation, high costs of production, low revenue generation, and uncertainties due to market dynamics among other issues [8].

In the context of the numerous challenges affecting stakeholders' efforts of ensuring food security, AI appears to be a tool that can help develop comprehensive food security management strategies, optimising crop yields, minimising losses, and reducing operational costs [9]. AI models offer significant advantages for food security purposes in terms of efficiency, accuracy, consistency, automation, pattern recognition, availability, and scalability, but they also come with challenges related to data, cost, ethics, and the need for ongoing management [10]. The choice between traditional methods and AI models depends on the specific problem, available resources, and the trade-offs involved. AI comprise analytical tools, such as data analytics, machine learning, and optimisation models, which can be applied to historical data and real-time information to make informed decisions [11]. This study focuses on the application of AI models. AI modelling is the key to building automated, intelligent, and smart systems ensuring that processes are better, faster, and more precise [12]. A combination of numerical optimisation, risk analyses, parameterization and scenario planning can help stakeholders address the multifaceted problem and adapt to changing conditions for long-term success [13].

The application of AI for food security involves a wide range of stakeholders and groups of people, such as farmers (e.g., commercial or smallholder farmers), agricultural researchers and scientists, food processing companies, food retailers and distributors, government agencies, NGOs, agriculture and food technology companies, and consumers [14]. Other stakeholders who also apply AI to enhance food security include investors and financial institutions, weather and climate agencies, supply chain and logistics companies, farm workers and labour forces, and food safety regulators, among other stakeholders [11]. The collaboration and coordination of these stakeholders are essential for addressing global food security challenges effectively [14,15].

In view of the stakeholders involved, in recent years, there has been a significant increase in investments by research agencies and governments worldwide in AI and its application in modelling food security issues [16]. Similarly, numerous research and policy organisations are actively developing models for effective policy making in the context of food security challenges [17]. Given the growing concerns regarding food (in)security, the scientific community recognises the crucial role of AI models in informing decision making. Despite the abundance of data and models available, there is often a failure in effectively translating these models into actionable policy [18–20]. A common characteristic of using AI in policy making is the interpretation of model outputs on a large scale [17,21,22]. While this approach is valuable, it is essential to integrate the local knowledge of stakeholders to ensure the robustness of model outcomes. The incorporation of stakeholders, their knowledge, and needs into models has been posited to result in better outcomes for stakeholders [23]. Involving stakeholders and incorporating their local knowledge in decision-making processes are necessary, as models alone can be ineffective tools [24,25]. Additionally, it is crucial to broaden the scope of model research to provide feedback to the communities under study [26]. Stakeholder participation requires collaboration with computer scientists and stakeholders [27], particularly the communities facing food insecurity. The approaches to stakeholder participation in AI modelling for food security vary [28].

Despite the diversity of participatory modelling methods to promote stakeholder involvement in modelling [29], particularly AI, the extent of stakeholder participation for food security purposes remains uncertain. Hence, there is little evidence that assesses stakeholder involvement in the application of AI models for food security. Also, limited knowledge exists on whether stakeholder involvement actually leads to a situation where AI modellers provide feedback to research communities. To this end, the main research questions that guide this study are as follows: What is the extent of stakeholder participation

in AI modelling for food security? How does this engagement influence the feedback loop between AI modellers and local communities facing food insecurity?

As AI gains prominence in the food security sector, several studies have been conducted on s indicators of the technology. Similarly, studies have been conducted on food security and modelling (see, for instance, [16,18,21,30–34]). However, the existing studies consist of fragmented themes and methodologies in the agrifood sector, characterised by diverse objectives across different disciplines, with relatively few connections between AI and food security. Recognising the knowledge gap, this study aims to analyse the AI models used in the literature for food security research

This study contributes to the growing importance of AI in enhancing food security by emphasising significance of collaboration between computer scientists and stakeholders to incorporate local knowledge into AI models, making them more effective tools for decision making.

2. Literature Review and Conceptual Framing of the Research

This section defines the key concepts such as food security and AI modelling applied in this study.

- Defining food security

A sufficient global food production alone does not guarantee food security for the entire population [35]. Hence, nutrition security complements food security by considering individuals' ability to meet their nutritional needs through food intake [36]. Nutrition security is typically assessed at the individual level, taking into account factors such as gender, wealth, age, and other determinants that affect an individual's access to food within households [37]. Despite these differences, food and nutrition security indicators are commonly aggregated at local, regional, national, and global levels for policy-making purposes [38]. Therefore, in this paper, the term "food security" encompasses both food and nutrition security across scales.

Four indicators of food security—food accessibility, availability, affordability, and utilisation—are operationalised for analysis in this study [39,40]. Access refers to the amount of food that can be produced, purchased from the market, or obtained through other means [36,37]. The production of food is primarily connected to the food availability indicator of food security [41]. Affordability, the third dimension, pertains to the cost or price aspect of food, while utilisation refers to households' ability to process accessible food. This depends on their capacity to acquire sufficient fuel, water, and other resources specific to their contexts. Utilisation also relates to individuals' physiological ability to digest food [42].

Given the complexity of the challenges hindering the attainment of the indicators of food security, the sector requires innovative solutions to mitigate the trade-offs between environmental, economic, and social objectives while ensuring the short- and long-term accessibility, availability, affordability, and utilisation of food [43,44].

- The application of AI models in the context of food security

Following Soori et al. [11], AI models can be categorised as follows: (i) machine learning models; (ii) neural network and deep learning models; (iii) data mining, knowledge discovery, and advanced analytics models; (iv) rule-based models and decision-making models; (v) fuzzy logic-based approaches; (vi) knowledge representation, uncertainty reasoning, and expert system models; (vii) case-based reasoning models; (viii) text mining and natural language processing models; (ix) visual analytics, computer vision, and pattern recognition models; and (x) hybridisation, search, and optimisation techniques. Hence, the AI models considered in this study encompass machine learning and deep learning models [27,45,46], including approaches that might not be typically considered as conventional AI; however, they have integrated AI techniques or algorithms to some extent in the models. AI models span a range from the global scale (e.g., food trade equilibrium

models) to the local scale (e.g., farm-level crop models, bioeconomic models, or agent-based approaches) [30].

The design or framework of AI models is typically top-down, neglecting the local setting, local knowledge, and the goals and needs of study communities. In such an instance, AI models utilise cadastral, biophysical, bioeconomic, or socioeconomic data or data integration, but they fail to incorporate the perspectives of the study communities. Critics argue that this approach results in detached outcomes that do not accurately reflect real-life situations and may not address actual food security issues. Müller et al. [17] pointed out that the exclusion of stakeholders and relevant concerns in these models reveals a gap that reflects social injustice.

Consequently, decisions often involve a trade-off between choosing the "right" model for the task and selecting the easiest model for researchers to use [3,47]. Therefore, modelling research is often driven by the capabilities of models rather than the requirements of stakeholders. Although some models explicitly aim to incorporate stakeholders' views, it remains challenging and unsatisfactory because predicting the future of food security under long-term global change scenarios is difficult. Additionally, modellers may encounter difficulties when combining different models due to the lack of empirical or experimental data for model parameterisation [48].

The complexity of global change and its interactions with the food system necessitate a multi-system and interdisciplinary perspective in AI modelling to address intricate food system issues. Consequently, an emerging approach, known as participatory modelling [49–51], seeks to integrate stakeholders' perspectives to bridge the existing gaps and overcome other challenges. More recently, analyses conducted at the global or regional level have made progress in finding solutions that better consider the local realities of individuals (e.g., [52,53]). Furthermore, recent advancements have witnessed collaborative efforts among modellers from diverse organisations, nations, and fields of expertise. This collective synergy aims to significantly strengthen worldwide food security [17].

3. Research Method

This study applied a literature review method [54]. The systematic review involved the identification, selection, and critical appraisal of relevant research, and an analysis of data from selected studies. To ensure a high level of quality reporting, this study adhered to the PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) guidelines [55]. This framework encompasses a checklist and a flow diagram, both of which are crucial for transparent reporting (see File S1 in Supplementary Materials). Figure 1 illustrates the flow diagram, providing a visual representation of the steps undertaken and briefly discussed in the following sub-sections.

3.1. Literature Identification

In the quest to identify studies pertaining to AI's role in food security, the study opted to use the Google search engine due to its comprehensive coverage of disciplines and interdisciplinary journal articles. The search queries included combinations of terms such as "Agricultural subsidy policies" AND "Food security" AND "Artificial Intelligence", "Farming systems" AND "Food security" AND "Artificial Intelligence", "Landownership" AND "Food security" AND "Artificial Intelligence", and "Cropping systems" AND "Food security" AND "Artificial Intelligence". The authors deliberately sought articles that focused on the concepts of food security and AI models (see Supplementary Materials File S1). While this approach aligns with the research objectives, it is worth noting that other studies and concepts may exist. Consequently, this article should be considered a foundational step in the discourse on the application of AI models for food security, offering insights for future research and collaborative efforts. The search was not limited to specific journals or subsets of the literature. Given this study's emphasis on AI modelling in the context of food security, the authors conducted a broad search using terms such as "Food security", "Artificial Intelligence", "Machine learning", and "Modelling" during the abstract screening

process, examining the title, abstract, and/or keywords of all articles in the database. This search was performed in December 2022 and yielded 1400 papers. To assess their relevance, the abstracts, introductions, and discussion sections, were thoroughly reviewed, resulting in the inclusion of 389 documents. Subsequently, 262 documents were selected due to removal of 127 duplicate documents.

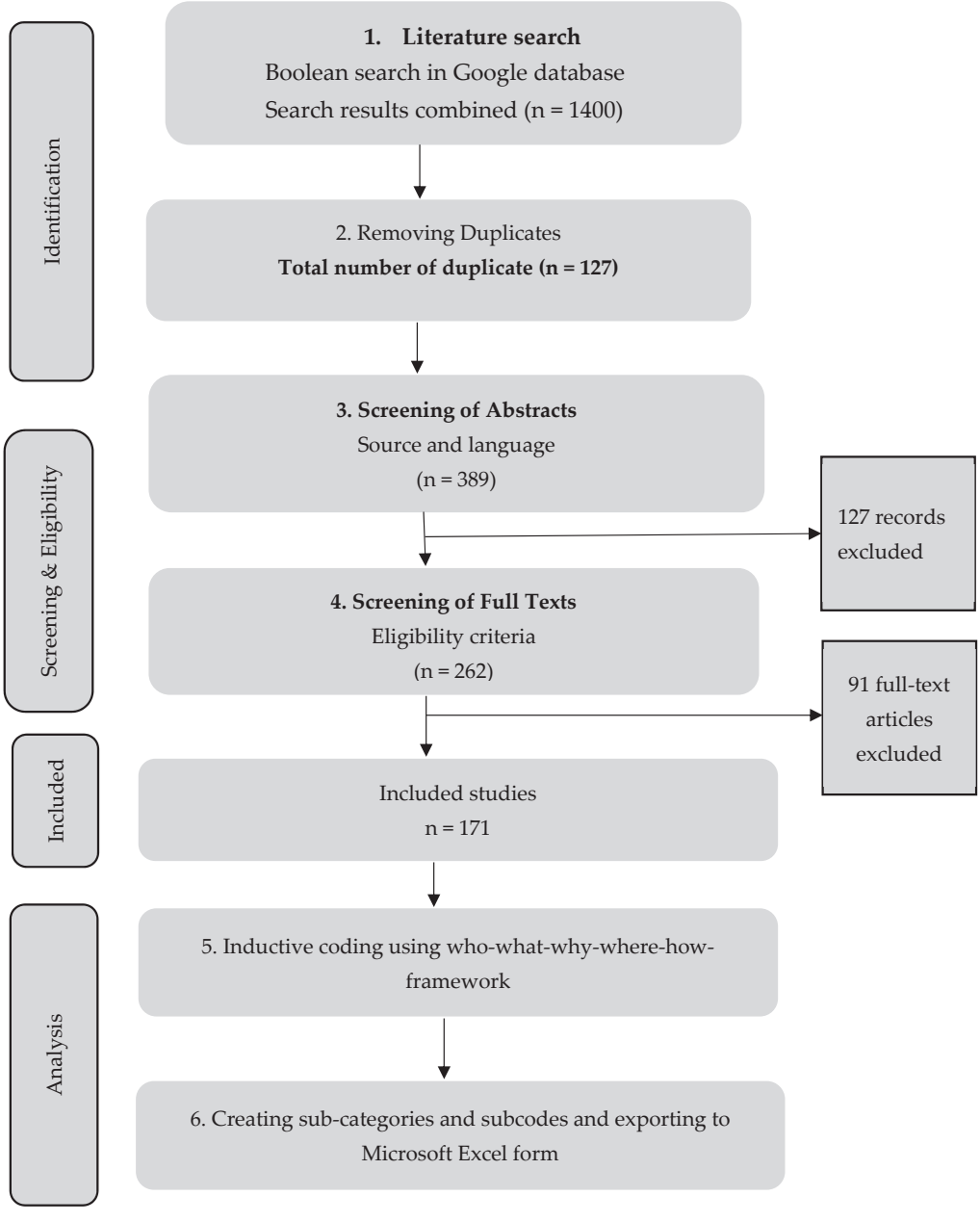


Figure 1. Schematic diagram of systematic review process.

3.2. Literature Screening and Eligibility

To assess the eligibility of the retrieved documents, this study established inclusion criteria. Studies failing to meet these criteria were excluded (n = 91). The selection criteria encompassed explicit references to methodologies for primary data collection and analyses concerning AI models or AI-related models, or indicators of food security. Another criterion focused on studies published in the English language from 2000 to 2022, as this period reflects the emergence of AI technologies and global food security discourses. Geographical coverage was also considered in our criteria. Ultimately, 171 articles met all inclusion criteria.

3.3. Analysis of Articles

The selected articles were imported into Atlas.ti software (version 22.1), where they were examined and categorised based on objectives, research methods, discussion, and conclusions. Atlas.ti's software capabilities facilitated the systematic organisation, coding, and analysis of qualitative data, such as text. It enabled annotation and data segmentation, the creation and application of hierarchical coding schemes, and the generation of reports and summaries. This software proved valuable for conducting in-depth content and thematic analyses of the selected documents. Each study underwent a thorough reading and coding process.

The method for analysing the selected documents primarily followed an inductive approach. In the initial phase, coding began with a list of codes for broad categories guided by the "who-what-where-why-how" concept. For example, under "where", coding initially encompassed geographical locations and subsequently generated sub-codes for specific regions (e.g., Asia, Africa, America, and Europe). For "why", coding examined the rationale, and open coding was employed to capture various aspects. Similarly, "how" aspect followed open coding, encompassing all relevant activities, including the models applied in each study. This approach allowed for the development of a comprehensive and detailed understanding of all the documents, linking broad themes and sub-codes together for the purpose of report writing. Coding encompassed the authors' institutions, funding organisations, study country/community, the model's end use, and food security indicators. Additionally, the theoretical underpinnings of the selected papers, focusing on the key concepts and constructs relevant to AI model research on food security (see Section 2), were coded.

Subsequently, codes for overarching themes (broad codes), such as geographical locations, indicators of food security, institutions conducting AI modelling research, and organisations funding AI models for food security research, were exported into Microsoft Office Excel 365 (Microsoft Corporation, Redmond, WA, USA). During this phase, more specific sub-codes were identified inductively. For instance, five sub-categories (i.e., development agencies, research institutions, state organisations, financial institutions, and others) were derived under the theme "organisations funding AI models for food security research". Subsequently, the data on categories and sub-categories in the extraction table were analysed for frequency of occurrence, leading to the creation of visual diagrams with Microsoft Office Excel. The results of the analyses are presented in the Findings Section. The analyses were not meant to generalise findings but rather identify pertinent research avenues for scholars to pursue context-specific investigations.

4. Findings

This section presents the findings categorised into sub-sections.

4.1. Overview of Geographical Locations of AI-Based Modelling Research on Food Security

The analysis of the selected documents showed the application of AI models for food security research in 68 countries across 6 continents. The distribution of the studies conducted in each continent is as follows: Africa (58 countries), Europe (17 countries), South America (16 countries), Asia (38 countries), North America (9 countries), and Australia

(4 countries). Additionally, some studies (30) had a global focus or focused on countries in different continents.

Among the 68 countries identified, approximately one-third of the studies were from the Global North [56–58], while two-thirds were from the Global South [24,59,60], with a particular emphasis on sub-Saharan African countries, like Ethiopia, Ghana, Malawi, Uganda, and Madagascar [61–67]. The European countries where research on AI and food security was conducted include Hungary, Spain, the UK, Italy, Poland, the Netherlands, Germany, and Greece [68–72]. Figure 2 highlights the geographical distribution of the identified countries and regions in the analysis.

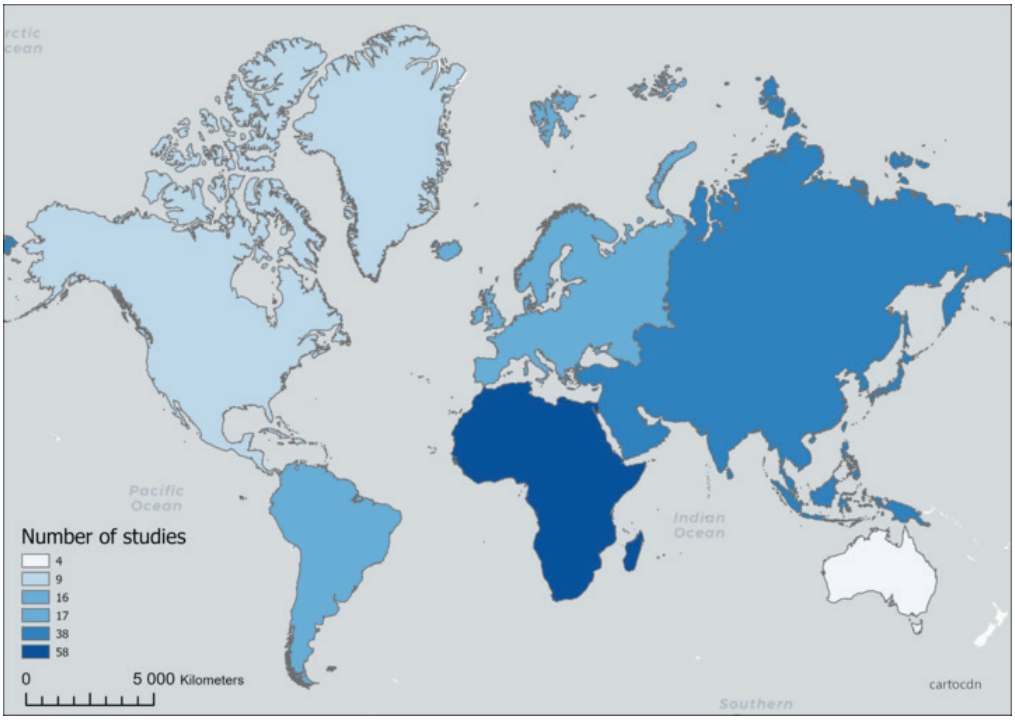


Figure 2. Geographical distribution of AI-based analytical research on food security. Source: authors' construct based on literature review, 2023.

4.1.1. Indicators of Food Security with the Application of AI Models

Research on AI and food security has analysed indicators, including accessibility, availability, affordability, and utilisation (Figure 3). However, most studies have primarily focused on the availability aspect of food security, often overlapping with other dimensions.

Regarding the accessibility indicator, 35 documents applied AI models to assess disparities in food distribution among households. These studies demonstrate how households rely on social safety programs or adopt coping strategies such as rural migration or the selling of livestock to improve their food security status [73–75].

Out of the 171 documents analysed, 105 documents mainly focused on food production, representing the availability indicator of food security [41]. The following themes provide an overview of the topics where AI has been applied to examine the availability indicator of food security:

- The impact of climate change on crop production, soil health, pest and disease patterns, and the vulnerability of crops [60,76,77].

- Predictions of changes in arable land cover, land use, and land use management under climate change [56,78–81].
- The effects of soil fertility decline, population growth, and poverty dynamics on food availability [59,82] and the application of climate or seasonal forecasts for food crop production [67].
- Agricultural productivity; crop yield predictions; risk; and the application of inputs, such as fertilisers, cropping patterns, sowing dates, crop variety selection, and cultural practices (like mulching and cropping patterns [43,65,79–87].
- Climate change and the contribution of livestock, aquaculture, and fisheries sectors to food availability [83,84].
- The potential of agricultural production systems, including smallholder farming systems, climate-smart agriculture, agroecology, agroforestry, industrial crops, organic farming, and the sustainable intensification of food crops [49,51,64,72,85–88].
- Farmers' behaviour, drivers, decision making regarding cropping patterns, input applications, and climate adaptation measures [89,90].
- Water availability, including groundwater and freshwater resources, irrigation dynamics, and management [25,91–94].

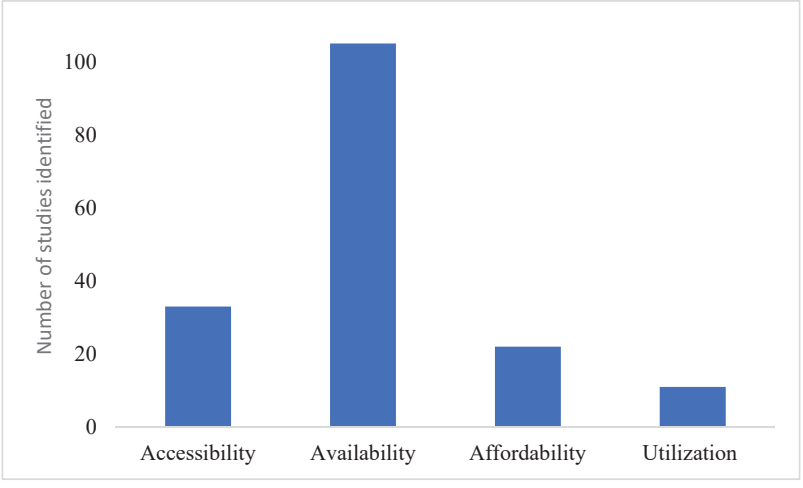


Figure 3. Indicators of food security analysed with AI-based model.

The crosscutting issues identified under the availability indicator of food security, with the application of AI models, include gender, labour issues, migratory household patterns [61,81,95–98], finance [99], cooperatives, local knowledge, infrastructure, supply/value chain [100–106], the energy–water–food nexus [107,108], and the application of digital technologies.

The affordability indicator of food security in AI model studies was limited to 22 documents, which explored the effect of subsidy policies on the food security of rural households [70,109]. These studies used models to create scenarios, examining conditions such as the taxation of groundwater extraction, increased water costs, reductions in agricultural subsidies, and cash transfer programs and the effects on rural livelihoods and food security [110]. Additionally, two document nvestigated the impact of food prices on food security and household income [62,111], particularly in sub-Saharan Africa. An AI model based on machine learning algorithms was also employed to estimate the role of scarcity, prices, food riots, and politically unstable regions [112].

AI models were also applied to assess the utilisation of food. The analysis also showed 11 documents focusing on social influence and the consumption patterns of specific food products, such as dairy [58,113,114] and plant-based meat alternatives. With this category,

models were also used to explore the food utilisation of different farming styles and their impact on hunger and rural health [115]. Other identified themes included social capital and the nutritional status of various groups, mainly rural farm households [116–118].

In this study, food security indicators were assessed at multiple levels, ranging from local to global (see Table 1).

Table 1. The application of AI models for food security indicators at various levels.

Indicators	Levels of Application of AI Models for Food Security Research				
	Local	**Sub-Regional**	**National**	**Regional**	**Global**
Accessibility	n = 23	n = 1	n = 8		n = 3
Availability	n = 47	n = 17	n = 11	n = 5	n = 23
Affordability	n = 19		n = 2		n = 1
Utilisation	n = 6		n = 1	n = 1	n = 3

Availability was the most frequently measured food security indicator across all levels, with data available at the local, sub-regional, national, regional, and global levels. This suggests a strong focus on assessing the quantity and availability of food. Accessibility indicators were examined at the local level (e.g., communities, villages, household, farm levels, and suburbs), in 23 documents. This assessment focused on whether households within these specific localities have sufficient access to food. The sub-regional level (e.g., watersheds, river basins, sub-regions, and provinces) was also considered in this study, with data collected from 1 document. This study extended its analysis on the accessibility indicator to the national level, involving data from 8 documents. The regional level (e.g., the European Union) analysis involved 3 documents, and, at the global level, this study examined food security indicators from a broader perspective. It included data from 23 documents.

Accessibility was also researched at various levels, including local, national, and regional levels, indicating an interest in understanding whether people have physical and economic access to food. Affordability was identified at the local and national levels, focusing on whether individuals and households can afford to purchase food. However, it was not assessed at the sub-regional, regional, or global levels. Utilisation was the least measured indicator, with data primarily available at the local and regional levels. Utilisation assesses how well individuals within a household utilise the available food, including nutritional aspects. It was not assessed at the sub-regional, national, or global levels. There was limited data at the sub-regional, regional, and global levels for most indicators, showing that food security assessment is often conducted at the local or national levels. The distribution of data across these indicators and levels suggests that food security assessments may vary in scope and depth, with a stronger emphasis on assessing the availability and, to a lesser extent, the accessibility and affordability of food. The absence of data at the global level for some indicators may be due to the challenges of collecting consistent with specific indicators of food security.

4.1.2. Model Types

This study identified various types of models; however, there were differences in their applications. Some studies applied a combination of models, including systemic and dynamic modelling. For instance, the ALUAM-AB model, an economic land use model based on Linear Programming Language (LPL) and a CPLEX solver, was applied to assess land use changes and their corresponding effects on food production and farming decision making. The models mentioned in the reviewed documents analysed human factors (such as the adoption of technologies, labour availability, and migration patterns), natural systems (weather, soil, water, etc.), or a combination of both. Machine learning algorithms, such as Boosted Regression Trees (BRT), Random Forest (RF), and Maximum Entropy (MAXENT) algorithms, were identified, mainly focused on analysing natural

systems. Additionally, agent-based modelling approaches, like the Common Resources Management Agent-Based System (CORMAS), Companion Modelling (ComMod), and RaMDry, were applied. As mentioned in Section 2, the identified models in Table 2 also include techniques that used AI algorithms at some stages.

Table 2. List of models applied in research on indicators of food security.

Models Applied in Food Security Research	Application in the Reviewed Documents	Type of Approach Adopted
Farm management model FARMACTOR and crop growth model system EXPERT-N	Crop management and decision making regarding planting and harvesting [69]	Natural systems
GIAM.GTEM [119]; Global and Local Learning Model, GALLM) (Hayward and Graham, 2013) [120]	Explored the response of land use and agricultural production to changes in productivity rates, resource scarcity and degradation, greenhouse gas abatement policy, climate change, and global demand [78]	Natural systems
Statistical Analogue Resampling (STAR) scheme, Weather and Research Forecasting (WRF) model, and Model for Nitrogen and Carbon in Agroecosystems (MONICA)	Evaluated the impact of two climate change scenarios on the profitability of double-cropping systems [60]	Natural and human systems
GIS modelling, Analytic Hierarchy Process (AHP), and an optimisation functionality	Assessed the "Energy-Water-Food nexus node" to support decision making for sustainable and resilient food security [108]	Natural systems
Markovian cellular automata and an agent-based approach	Investigated the future land use trajectories of a semi-arid Mediterranean agroecosystem [121]	Natural and human systems
Computable General Equilibrium (CGE) model	Seasonal rural labour markets and their relevance to policy analyses [81]	Natural and human systems
Discrete Event Simulation	Simulated potential growth strategies and observe the impact concerning existing farm processes [122]	Natural and human systems
Change detection methods and agent-based modelling	Examined of historical and future land use changes [118]	Natural and human systems
ADOPT (combines socio-hydrological and agent-based modelling approaches by coupling the FAO crop model AquacropOS with a behavioural model capable of simulating different adaptive behavioural theories)	Evaluated the factors that influence adaptation decisions and the subsequent adoption of measures and how this affects drought risk for agricultural production [90] Farmers facing droughts: capturing adaptation dynamics [25] Education, financial aid, and awareness can reduce smallholder farmers' vulnerability to drought [99]	Natural and human systems
The ID3 rule induction/machine learning algorithm	Assessed farmers' adaptation to changes in environmental and economic contexts [68]	Natural and human systems
MOSAICA	Assessed the upscale of CSA [51]	Human systems
Machine learning algorithms, including Boosted Regression Trees (BRT), Random Forest (RF), and Maximum Entropy (MAXENT) algorithms	Mapped the suitability for small-scale, informal irrigation [94]	Natural systems
GIS and agent-based modelling	The interlinkage and interaction of resource–food–bioenergy systems and optimise supply chains considering poly-centric decision spaces [123]	Natural systems
Remote sensing and Artificial Intelligence techniques (neural network algorithms)	Identified of food insecure zones [95]	Natural systems

Table 2. *Cont.*

Models Applied in Food Security Research	Application in the Reviewed Documents	Type of Approach Adopted
CLASSES (bioeconomic system dynamics model) and Mexico Sheep Sector Model (MSSM)	Illustrated how three indicators of access (food consumption expenditures, a food insecurity scale, and dietary diversity) and their stability can be incorporated into a dynamic household-level model of a maize-based production system and a dynamic regional model of sheep production and marketing [98]	Natural and human systems
Artificial Intelligence and deep learning approaches	Agricultural productivity and crop yield predictions and risk [124–130]	Natural systems
Model for Nitrogen and Carbon Dynamics in Agroecosystems (MONICA) and Mathematical-Programming-based Multi-Agent Systems (MPMASs)	Identified biophysical and socioeconomic dimensions of yield gaps [59]	Natural and human system
MPMASs with the crop growth model Model for Nitrogen and Carbon in Agroecosystems)	Examined farmers' decision making and agricultural land use to account for the interplay between the environment and human decision making [89]	Natural and human system
Common Resources Management Agent-Based System (CORMAS) and SimSahel model	Tested the impact of social forces on the evolution of Sahelian farming systems [85]	Natural and human systems
Asia-Pacific Integrated Model (AIM) using Computable General Equilibrium	Inclusive climate change mitigation and food security policy [96,97]	Natural and human systems
Agent-based model	The potential effects of a subsidised policy on households to rent out land use rights for long terms under formal contracts and impacts on food security [109]	Natural and human systems
	Examined impacts of climate and price variability on household income and food security [62]	Natural and human systems
	Assessed of household-level and community-wide resilience to climate shocks in a smallholder mixed crop–livestock farming setting [84]	Natural and human systems
	Assess future patterns of arable land use under four localised, stakeholder-driven scenarios of plausible future socioeconomic and climate change [80]	Human systems
	Integrated of seasonal precipitation forecast information into local-level agricultural decision making [67]	Natural and human systems
	Milk consumption and scenarios of dairy reduction and adoption of plant-based milk (PBM) [58]	Natural and human systems
	Simulated the impacts of climate variability and change on crop varietal diversity [131]	Human systems
	Irrigation agriculture dynamics [91,93]	Natural and human systems
	Explored how interactions between households and the environment lead to the emergence of community food availability, access, utilisation, and stability over time [61]	Natural and human systems
	The impacts of cash transfer programs on rural livelihoods [110]	Natural and human systems
	Transition from conventional to organic farming [91]	Natural and human systems

Table 2. *Cont.*

Models Applied in Food Security Research	Application in the Reviewed Documents	Type of Approach Adopted
Agent-based model	Assessment of agricultural vulnerability of sugarcane facing climatic change [77]	Natural and human systems
	Characterising farm types and evolvement in smallholder dairy systems [113]	Natural and human systems
	Assessed the impacts of the changes in farming systems on food security and environmental sustainability of a rural region [76]	Natural and human systems
	The management of aquaculture production [83]	Natural and human systems
	Vulnerable households using migration to manage the risk of rainfall variability and food insecurity [74]	Natural and human systems
	Market of potato (Solanum tuberosum) producers [100]	Human systems
	The impact that water canals and electric grid development have on the Water–Energy–Food (WEF) nexus in a rural area [107]	Human systems
	Analysed the impact of climate-smart agriculture on food security using an agent-based analysis [88]	Human systems
	Simulate strategies of the perishable food market under different circumstances [104]	Human systems
	Reduced meat consumption [114]; social influence on meat-eating behaviour [132]	Human systems
	Examined the supply chain of organic fertiliser [102]	Human systems
	Evaluated food supply chain resilience: potato supply chain [105]; contract farming in rice supply chain [103,106]	Human systems
	Examined the uptake of new farming practices, for example, organic waste application [87]	Human systems
	Examined household food security, climate outlook, and agricultural productivity [133–137]	Natural and Human systems
	Examined climate change, hunger, and rural health through the lens of farming styles [115]	Natural and human systems
	Simulated small-scale farmers' agroforestry adoption decisions to investigate the consequences for livelihoods and the environment over time [138]	Natural and human systems
	Projected the effect of crop yield increases, dietary change, and different price scenarios on land use under two different state security regimes [111]	Natural and human systems
	Examined Food security and global trade [101]	Human systems
	Asses the potential of land use change for mitigation of food deserts [56]	Natural and human systems
	Analysed the diffusion of added-value markets among Dutch farmers [139]	Natural and human systems
	Examined disparities in food accessibility among households [117]	Natural and human systems
	Farmers' adaptation to agricultural risks [140]; adaptive management in crop pest control in the face of climate variability [141] postharvest loss of food grains [142]	Natural and human systems

Table 2. *Cont.*

Models Applied in Food Security Research	Application in the Reviewed Documents	Type of Approach Adopted
Dimensions of Agent-Based Modelling Approaches		
Common Resources Management Agent-Based System (CORMAS) [143,144]	Assessed development intervention on the provision of fertiliser and credit to farmers [73]	Natural and human systems
TERROIR (TERRoir level Organic matter Interactions and Recycling model)	Analysed nutrient cycles at three levels of organisation: plot, household, and landscape [145]	Natural and human systems
AMBAWA model	Assessed the impacts of the practice of crop residue mulching on crop productivity [64]	Natural and human systems
Spatially explicit empirical agent-based model (SEALM)	Examined possible future trends of farmers' crop management and the effects of these trends on the environment, household economy, food self-sufficiency, and household coping strategies for food insecurity [65]	Natural and human systems
Companion Modelling (ComMod)	Assessed the synergies and trade-offs between REDD+ and climate-smart agriculture [49] Groundwater irrigation management with local farmers [92]	Natural and human systems
ALUAM-AB (an economic land use model based on Linear Programming Language (LPL) and a CPLEX solver)	Assessed the interaction effects of these agricultural policies while accounting for climate change impacts in the analysis [70]	Natural and human systems
Integrated assessment modelling (IAM) using coupled component modelling (CCM) approach to derive an agent-based model associated with a soil model and multi-scale spatial model, resulting in the Model for West-Africa Agroecosystem Integrated Assessment (MOWASIA)	Assessed the environmental and economic performances of semi-continuous and continuous farming systems [146]	Natural and human systems
Agent-based rangeland model RaMDry	Assessed the vulnerability of rangelands and livestock production systems as a result of the effects of ongoing changes in precipitation and its variation, as well as its temporal distribution [75]	Natural systems
Multi-agent systems (MAS)	Simulated soil fertility and poverty dynamics [147] Simulation of the sustainability of farming systems [148]	Natural and human systems
Mathematical Programming-based Multi-Agent Systems (MPMASs)	Analysed how adaptation affects the distribution of household food security and poverty under the current climate and price variability [82] Analysed of the biophysical and socioeconomic factors that influence the livelihood strategies of traditional Andean farmers and study how these systems are being affected by climate change [149] Climate variability, social capital, and food security [143] Examined watershed-level irrigation management [150]	Human systems
The OMOLAND-CA (OMOLAND Climate Change Adaptation) model	The socio-cognitive behaviour of rural households towards climate change and resource flows prompt agents to diversify their production strategy under different climatic conditions [116]	Models natural and human systems

Table 2. *Cont.*

Models Applied in Food Security Research	Application in the Reviewed Documents	Type of Approach Adopted
The farm management model (FarmActor)	Examined how climatic changes drive farmers' adaptation of their land use decisions [151].	Models natural and human systems
Common Resources Management Agent-Based System (CORMAS)	Analysed the impact of development interventions on the rural population [85]	Models natural and human systems
Flows in Agro-Food Networks (FAN)	Simulated contrasting scenarios of material flows locally in a small farming region [152]	Models natural and human systems
The Dawe Global Security Model	Simulated the global food market, food riot, and the political fragility of countries [112]	Human systems

4.2. Institutions Conducting AI-Based Modelling Research for Food Security

This study identified three categories of institutions that engaged in AI modelling research for food security. These categories are as follows:

1. Local organisations based in countries where the research was conducted, predominantly comprising universities or research institutions [70,76,78,108,122,131,153].
2. Collaboration was observed between local and foreign research institutions [59,60,80,81,113,121,154].
3. Foreign organisations, including universities and international research organisations, solely focused on AI research related to food security [24,49,62,65,66,73,84,109].

According to Figure 4, AI modelling research on food security was primarily led by foreign organisations. These collaborations involved universities and research institutions from the Global North, while the actual study communities were based in the Global South. The local organisations engaged in AI modelling research for food security in home-based countries were predominantly from developed countries [69,70,78]. However, there were exceptions, as local organisations carried out research in countries such as Qatar, Iran, China, South Africa, and Saudi Arabia [70,92,104,108,155].

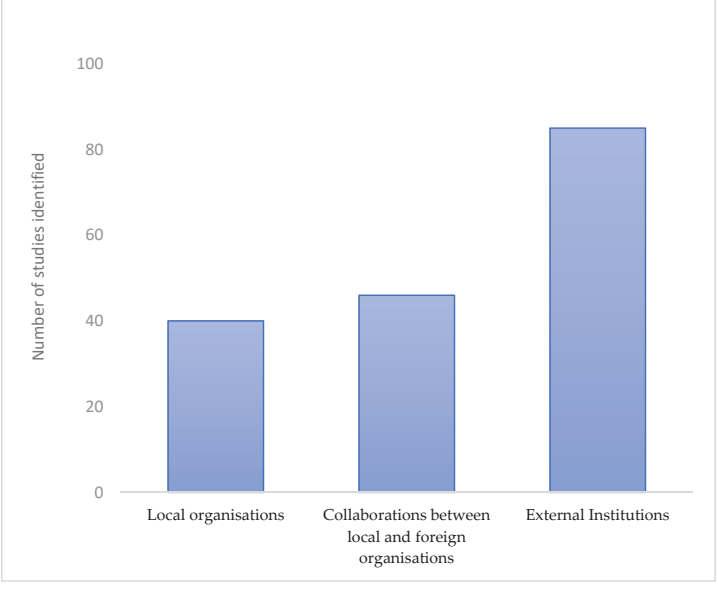


Figure 4. Origin of Institutions conducting AI-based modelling research on food security. Source: authors' construct based on literature review, 2023.

4.3. Categories of Organisations Funding AI Modelling Research on Food Security

This study identified three categories of funding for AI modelling research on food security as follows:

i. Foreign funding: AI modelling research on food security was funded by development agencies, research institutions, state organisations, financial institutions, and other forums (e.g., foundations or platforms) [24,31,49,59,64–66,84].
ii. Collaborative funding: Funding between local universities, research institutions, and foreign partners [89,155].
iii. Local funding: Home-based institutions (e.g., local universities and research institutes) received funds from governments through research councils [58,69,70,72,78,108,113].

Figure 5 illustrates findings on the three modes of funding for AI modelling research for food security. However, collaborative funding (i.e., the contribution of funds between a home-based institution and external partners) was less than sole foreign funding. In some instances, home-based institutions provided support to research in kind. Of the 171 papers reviewed, 36 funding sources could not be identified, as the documents did not mention any funding organisations.

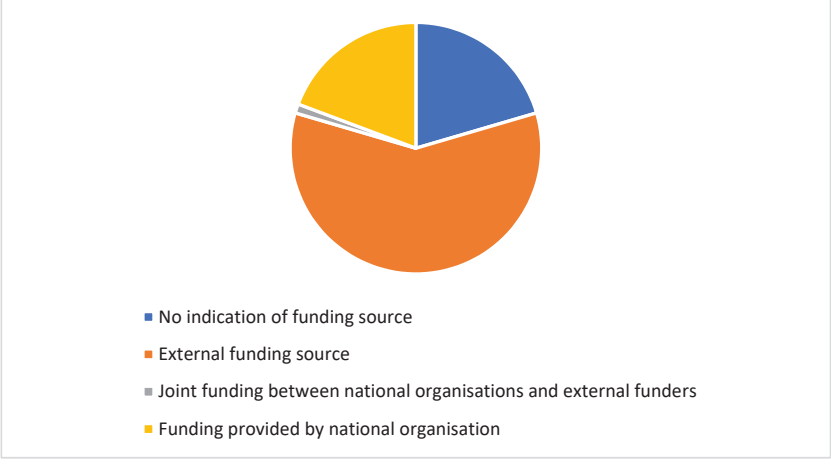


Figure 5. Sources of funding for AI-based modelling research on food security. Source: authors' construct based on literature review, 2023.

External funding for AI modelling research for food security was provided by development agencies, research institutions, state organisations, financial institutions, and others (refer to Figure 5). For example, funding for AI modelling research on food security was received through initiatives such as the CGIAR Program on Water & Food; "Meeting Urban Food Needs", a project by the Food and Agriculture Organization of the United Nations; the Global Hunger and Food Security Initiative of the United States Agency for International Development (USAID); and the West African Science Service Center on Climate Change and Adapted Land Use (WASCAL), among other initiatives (see Supplementary Materials File S2). Figure 6 illustrates the distribution of external funding sources for AI modelling research on food security. The analysis showed 103 reviewed documents that mentioned external funding. However, the analyses showed 150 organisations because some documents stated more than one external funding organisation.

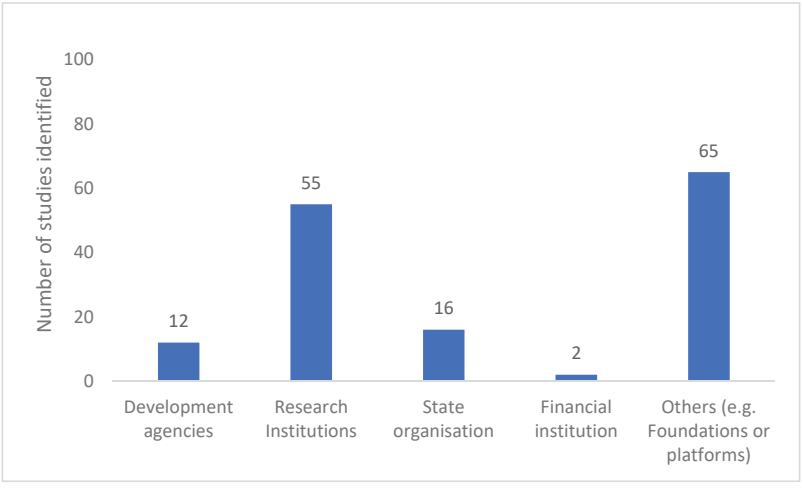


Figure 6. Sources of external funding for AI-based model research. Source: authors' construct based on literature review, 2023.

External funding for AI modelling research on food security has primarily focused on countries in the Global South, particularly Africa. These countries include Niger, Senegal, Ethiopia, Burkina Faso, Tanzania, Madagascar, Ghana, Uganda, Malawi, Kenya, Cameroon, and South Africa [61–67]. A similar funding pattern was observed in the Asian region, specifically in Laos, Vietnam, Bhutan, Thailand, Indonesia, India, Sri Lanka, the Philippines, Bangladesh, and Nepal [49,81,109,149]. In South America, the countries identified included Paraguay, Brazil, Colombia, Peru, Mexico, Chile, Guadeloupe, Uruguay, Ecuador, and Guatemala [24,59,60,91,100,149,156].

In the Global North, government and national/local organisations provided funding for AI modelling research on food security in countries such as Germany, Switzerland, the UK, Norway, Australia, the USA, Canada, the Netherlands, France, and Belgium [68–72]. In the Mediterranean and Arab (MENA) region, countries like Qatar, Iran, and Saudi Arabia have foundations or government-sponsored initiatives for AI modelling research on food security [108,155]. While most modelling studies on food security in the Global South were externally funded, there were a few exceptions where some local institutions also financed research activities on food security [76,136,157].

4.4. Approaches Used in AI Models for Food Security Research

This section analyses three main approaches commonly adopted in AI model research on food security. These approaches are as follows:

4.4.1. The Application of Only Artificial Intelligence Models for Food Security Situations

The first approach identified in the studies (87 documents) used various types of data without involving stakeholders. The data sources included biophysical data (such as crop data, weather/climate data, and soil characteristics data), bioeconomic data (including historical agricultural production data), and cadastral data (such as satellite images, geospatial information, and land use and landcover maps). In some studies, different datasets were integrated into models to investigate specific aspects of food security.

The recommendations and conclusions reflected on the model building process and the potential of AI to conduct specific analyses, frameworks, and methodologies. Ultimately, they offered insights into the possibilities and limitations of the models, such as the unavailability of data for model development, evaluation, and application. Furthermore, the studies drew conclusions on the future operationalisation of the model for food security and proposed avenues for future research, including the incorporation of param-

eters, variables, socioeconomic and environmental factors into the models. However, a few studies recommended involving stakeholders when applying the model outputs in real-life situations.

4.4.2. Involving Stakeholders in Artificial Intelligence Model Research for Food Security

The second approach identified in the reviewed literature (81 documents) involved integrating AI models, biophysical data, and primary data derived from stakeholders. Data utilised for this approach encompassed preliminary data gathered through various methods, including participatory research, surveys/questionnaires (using the Likert scale, open-ended formats, and close-ended formats), focus group discussions, interviews, Bayesian network analysis, workshops, role-playing games, community meetings, observations, and ethnographic approaches (refer to see Supplementary Materials File S3 for more details). In certain studies, mixed methods were employed to harness the benefits of integrating real-life data with models. Some studies utilized a research methodology involving field investigations, local experts' discussions, and interdisciplinary modelling. Some studies developed parameters and scenarios with stakeholders. Stakeholders also supported the modelling process, for example, participating in the validation and simulation stages. Additionally, studies identified in this category also applied secondary data in the form of a literature review, consisting of peer-reviewed literature, project reviews, and government documents outlining government directives, as well as food security strategies and planning. The secondary data sources (e.g., census data) were applied in some studies to construct scenarios for the models.

With this approach, models also applied well-known concepts from sociology, science communication, and economics to study indicators of food security. Although the studies reviewed under this category included stakeholders' perspectives, their main focus was building a modelling framework to explore the extent of food security indicators. The discussions of findings, recommendations, and conclusions proposed how policy and research results could enhance productivity and food security. Most studies analyzed the effectiveness of the methodology applied and recommended the use of mixed method approach and using an interactive and interdisciplinary approach to understand food security situations. Some studies in this category also recommended cross-validation and suggested pathways for future research, such as incorporating a larger sample size, developing the concept of value co-creation, and assessing the feasibility of different paths in practice.

4.4.3. Stakeholder Involvement in AI Modelling for Food Security through an Iterative Process

There was a limited application (3 documents) of this approach where stakeholders were involved in AI modelling for food security through an iterative process. Findings on this approach indicate the utilization of biophysical and real-life data, as described in Section 4.4.2. However, researchers shared their findings (feedback) with study communities and assisted in implementing the results. The modelling process involved workshops, conceptualization, scenario building, implementation, validation, and experimentation. Various methodologies combined tools and interpretations of remote sensing data in computer modelling (ComMod and multi-agent model). Stakeholder engagements involved conventional on-farm research, role plays, on-farm surveys, innovation platforms, farmer field schools, and experiments. The studies that applied these approaches reflected on the "best practice" of involving stakeholders, and it provided entry points to engage stakeholders, integrated knowledge from diverse sources and identified methodological choices that could be replicated. They also mentioned capacity-building initiatives, as well as monitoring and evaluation as means of critical reflections and learning on the successes and failures. Studies identified for the analysis under this category communicated or provided feedback to research communities. Conclusions also demonstrated that local

stakeholders' involvement enhanced the design of local interventions, tailoring them to fit the specific context.

5. Discussion

In this section, the key themes that emerged from the research findings are discussed.

This review conducted on the use of AI models for food security purposes revealed few studies that have demonstrated how feedback is incorporated into policy or even implemented the outcomes with stakeholders. The approach to go beyond AI modelling and stakeholder consultation was observed in three studies in the analysis that employed participatory modelling, involving stakeholders' knowledge to develop formalised and shared representations of reality. These studies used models as boundary objects to facilitate two-way learning and collective reasoning about food security challenges iteratively. The limited provision of feedback and implementation of outcomes in study communities, as identified in several studies, may be attributed to the current organisation of research, where there is more focus on model outputs for publications to fulfil aspects of funding indicators. Moreover, the limited timeframe of consortiums can result in a lack of continuity in applying research outcomes in communities. The funding issues and the organisations responsible for implementing AI modelling research may also influence the selection of research designs, including the provision of feedback and the implementation of findings.

This study also found that the sources of funding for AI modelling research on food security were predominantly external to study communities. Most of the reviewed studies indicated that researchers initiated the research process rather than the local population. This finding has implications for the implementation of AI models for food security to address local needs. Hence, it highlights the need for local stakeholders to take the lead in setting the research agenda, rather than relying solely on researchers and funding agencies. The findings also emphasises the importance of promoting the use of AI technologies for demand-driven innovation and scaling approaches.

The findings on the themes of food security, for example, climate-smart agricultural practices, agroecology, climate information services for agriculture, and sustainable agricultural practices, among other related themes, align with the current agenda set by international donor agencies and organisations for global food security and food systems. According to Verburg et al. [19], most research grants for modelling in the crop and agricultural sector, for instance, are driven by specific issues, with model development and implementation often serving as secondary objectives. This study also calls for the international community to critically review the approach to funding research, especially for AI modelling research on food security, as it has shed light on the missing link between researchers and stakeholders, especially local communities.

Also, studies did not adequately address or even document how they dealt with issues regarding inclusiveness and the equitable and fair representation of stakeholders and vulnerable groups in the selection of stakeholders. In addition, the documents analysed lacked explicit statements justifying the selection criteria of stakeholders. The documents reviewed also did not illustrate how they created separate interaction spaces for different groups of stakeholders. Additionally, the studies did not consider selection criteria for participating stakeholders, including researchers. The consideration of AI researchers in the food security sector as stakeholders in the research can strengthen the research framework and outcome for practical implementation.

The findings in Section 4.4.2 show that some studies used already established interaction spaces from previous or ongoing projects, such as innovation platforms and farmer field schools. These spaces incorporated stakeholders who were already actively involved in knowledge production, sharing, and use. Examples of such stakeholders include opinion leaders, extension agents with good relationships with farmers, experts recognised by their peers, and individuals with extensive networks. However, only a few studies explicitly stated how they applied interactive principles, such as involving stakeholders throughout

the process, creating opportunities for mutual learning, and monitoring and assessing mutual learning (as mentioned in Section 4.4.3).

While many models fall short in addressing inclusivity and fairness when dealing with food security challenges, establishing some points of departure can help provide context and relevance to modelling approaches. One such point is determining whether AI modelling approaches can solve all food security challenges. Answering this question requires selecting approaches that promote inclusivity and accurately reflect people's realities. However, considering the wide range of realities and priorities, achieving this goal may prove challenging, particularly for local communities. Moreover, modellers often tend to adhere to their conventional approaches and define realities based on their own preferences.

In the documents reviewed, especially those that involved stakeholders, there was an observed discrepancy where the studies tried to encompass broad perspectives while simultaneously addressing contextual issues. In this instance, there are risks associated with applying AI models to study food security on a global scale. However, there is also recognition regarding the issue of scale and the trade-offs associated with it, as it can influence the extent to which AI modelling research can facilitate social justice and equity in addressing real-life food security challenges. Consequently, these intriguing dilemmas prompt the need to consider how we can navigate the issue of scale in models.

In some studies, the primary goals of the models were focused on model building and framework development, rather than employing the models to solve specific food security challenges. Also, the development of models has been characterised by gradual, often ad hoc improvements and extensions, with different groups incorporating "new" modules that have already been implemented elsewhere [17]. Consequently, there has been an increase in the number of models, but insufficient attention has been given to other crucial processes related to food security, such as addressing perennial challenges, and bridging the gap between simulated and actual on-farm yields [19]. Here, this study proposes that, instead of viewing models as black boxes, they should be seen as internally logical tools that can lead to specific outcomes on food security challenges, thereby persuading stakeholders to adopt appropriate actions.

Contributions of This Study to Policy and the Scientific Community

One key contribution of this study is the identification of different categories of institutions involved in AI modelling research on food security. This study highlights the dominance of foreign organisations, mainly from the Global North, in spearheading AI modelling research on food security, despite the research communities being based in the Global South. This information is vital for policymakers and researchers to ensure inclusivity and equity in AI modelling research on food security, promoting a balance between local and foreign involvement. Another contribution of this study is the comprehensive overview of the organisations funding AI modelling research on food security. This article identifies three modes of funding for AI models, including foreign funding, collaborative funding, and local funding. Additionally, it presents a detailed analysis of external funding sources for AI modelling research on food security, with a focus on countries in the Global South, North, and MENA regions. By providing this information, this article offers valuable insights into the global funding landscape for AI modelling research on food security, which can inform future research and policy decisions in this field.

This study also identified three approaches: one focusing on biophysical and bioeconomic data and the other focusing on integrating local knowledge and experiences through stakeholder involvement. This study emphasises the importance of stakeholder participation in developing policy instruments to improve food security, such as subsidies, infrastructure deployment, economic and market incentives, and extension and advisory systems. Additionally, it discusses the limited involvement of stakeholders throughout the research process, including the provision of feedback and implementation of research outcomes, providing insights into the potential benefits of increased stakeholder engagement,

such as promoting participatory monitoring and evaluation and tailoring interventions to fit the local context.

The findings of the study have implications for policymakers, researchers, and practitioners working towards achieving food security and promoting sustainable development. Overall, this study contributes to the growing body of knowledge on the application of AI in addressing global challenges like food insecurity.

6. Conclusions

In conclusion, the analysis of geographical locations, indicators of food security, model types, institutions involved, and funding sources in AI modelling research on food security has revealed both the global scope and the intricate web of stakeholders engaged in addressing this critical issue. However, the multifaceted problems surrounding food security persist, demanding continued attention and innovative approaches for involving stakeholders. One of the key takeaways from this review research is the significant emphasis placed on the availability indicator of food security through AI modelling, with research predominantly concentrated in sub-Saharan Africa. While this focus is crucial, it is important to recognise that the affordability and utilisation aspects of food security also warrant closer examination, as they are interconnected with the broader food security landscape. Moreover, the integration of stakeholders, local knowledge, and community participation in AI modelling research on food security has demonstrated its potential for generating meaningful insights and real-life interventions. This approach can lead to a shift in community perceptions, promote trust, and empower marginalised groups to participate in decision-making processes. As we navigate the complexities of ensuring food security in a changing world, it is evident that collaborative efforts between researchers, stakeholders, and local communities are essential. Only by engaging in iterative processes that bridge the gap between theory and practice can we hope to address the multifaceted challenges of food security effectively. Through such inclusive approaches, we can drive meaningful change, enhance food security, and ultimately contribute to a more sustainable and equitable future.

Based on the findings of this study, this study proposes the following recommendations to enhance research on the application of AI and food security.

There is a need to build trust between researchers and study communities to maintain long-term collaborations. This can be achieved through informal, shared experiences that build social capital among researchers, stakeholders, and communities facing food security. Trust can be built when AI modellers (research groups) acknowledge, examine, and appreciate differences in the study community's traditions and experiences and select methodological approaches that foster engagement. It is essential for researchers to fully understand the scope of the study community and consider contextual factors that can affect decision making.

Computer science collaborators (i.e., AI modellers) should consider themselves as part of the food security project rather than just technical support or information technology consultants. They can also build a common vocabulary while explicitly exploring food security problems to help understand the contexts of study communities.

Researchers may create a culture of listening and open communication with stakeholders and ensure that modelling research provides the results to study communities. A lack of familiarity with research goals and modelling approaches may lead to a misunderstanding of the constraints associated with the application of AI models for food security. Therefore, the organisation of frequent meetings between modellers, researchers, and study communities can help overcome language barriers and provide real-time clarification of jargon or misunderstandings.

AI modellers can also engage with study communities early in the research design process to understand local needs and explain any limitations that cannot be met in the modelling research.

The provision of training and capacity building to researchers from different disciplines and stakeholders on basic quantitative skills, computational literacy, and scenario building can aid in the interpretation of research results, integration into policy, and other aspects of the research process. This can help study communities better understand model outcomes and enable modellers/researchers to communicate research goals for improved collaboration and communication.

Researchers can also help stakeholders develop Decision Support System (DSS) tools to enable feedback provision and advisories to communities facing food security challenges. The communication of the research outcome to stakeholders and study communities can also be carried out through various knowledge products, such as computer-based data visualisation tools or by linking modelling outcomes to mobile devices and apps.

Research funders should also ensure the provision of feedback and/or the implementation of findings to stakeholders or study communities as part of the funding requirement. Assessments of research proposals requiring funding should integrate the aforementioned criteria as components. The monitoring and evaluation of research should also ensure that researchers are in contact with stakeholders and study communities. Funding mechanisms are needed to support integration, continuity, shared learning, and continuous innovation.

Future Research

While this study provides insights into where and how AI models are applied for food security, there is a limited assessment of the extent (measurable) that the use of Artificial Intelligence helps or can help achieve the required level of food security. The identified limitations of this research are mainly attributed to the literature review approach adopted. Future research should focus on evaluating the continued effectiveness of AI models in improving food availability, accessibility, affordability, and utilisation over extended periods. As food security is influenced by evolving factors, such as climate change, socioeconomic shifts, and global crises (e.g., pandemics), future research should assess how AI models adapt to changing conditions and contribute to the resilience of food systems. Research should delve deeper into the participatory processes involving stakeholders in AI modelling. Understanding how local communities are empowered to make informed decisions and implement AI-driven recommendations can provide valuable insights. Given the diverse geographical locations where AI models are applied, future research should explore the cultural, contextual, and governance factors that influence the success or challenges of AI interventions in different regions. Comparative studies can also be conducted across countries and regions to identify the best practices, the lessons learned, and transferable AI-based approaches for enhancing food security. Future research should also investigate the ethical implications of AI models in food security that involve stakeholders, especially issues related to data privacy, fairness, and the equitable distribution of benefits. Research should also assess the potential unintended consequences of AI-driven decisions on vulnerable populations, including examining the role of policies, regulations, and governance structures in supporting or hindering the implementation of AI models in food security initiatives. Future research should also evaluate the scalability and replicability of successful AI-driven interventions, considering the feasibility of adapting these models to diverse food security contexts. Addressing these research gaps will provide a more holistic understanding of the role of AI in sustainable food security and inform the development of effective, context-specific strategies for enhancing food security worldwide.

Supplementary Materials: The following supporting information can be downloaded at: https://www.mdpi.com/article/10.3390/agriculture13102037/s1, File S1: Query search applied in the google search engine; File S2: List of funding organisations that sponsored AI-based analytical research; File S3: Type of data applied in approach two.

Author Contributions: R.S., U.A.C. and T.C.: conception of the research and outline of the paper. R.S.: methodology, writing—original draft, writing—review and editing, and visualisation. U.A.C.: resources supervision, crosschecking of selected documents, and writing—review and editing. T.C.: conception of research, methodology, and writing—review and editing. All authors have read and agreed to the published version of the manuscript.

Funding: This research was conducted under the sponsorship of the Alexander von Humboldt fellowship. However, the views expressed herein are the sole responsibility of the authors.

Acknowledgments: We thank Linda Petutschnig in the Multi-Agent Research Group, Hamburg University of Applied Sciences, for supporting the creation of the map in our research analysis. We acknowledge support for the article processing charge by the Open Access Publication Fund of Hamburg University of Applied Sciences.

Conflicts of Interest: The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

References

1. Sarker, I.H. Machine Learning: Algorithms, Real-World Applications and Research Directions. *SN Comput. Sci.* **2021**, *2*, 160. [CrossRef] [PubMed]
2. Bhagat, P.R.; Naz, F.; Magda, R. Artificial intelligence solutions enabling sustainable agriculture: A bibliometric analysis. *PLoS ONE* **2022**, *17*, e0268989. [CrossRef]
3. Vinuesa, R.; Azizpour, H.; Leite, I.; Balaam, M.; Dignum, V.; Domisch, S.; Felländer, A.; Langhans, S.D.; Tegmark, M.; Nerini, F.F. The role of artificial intelligence in achieving the Sustainable Development Goals. *Nat. Commun.* **2020**, *11*, 233. [CrossRef]
4. *The State of Food Security and Nutrition in the World 2022*; FAO: Rome, Italy, 2022.
5. van Dijk, M.; Morley, T.; Rau, M.L.; Saghai, Y. A meta-analysis of projected global food demand and population at risk of hunger for the period 2010–2050. *Nat. Food* **2021**, *2*, 494–501. [CrossRef]
6. Ritchie, H. Smallholders Produce One-Third of the World's Food, Less than Half of What Many Headlines Claim. Our World in Data. 2021. Available online: https://ourworldindata.org/smallholder-food-production (accessed on 26 September 2023).
7. Hebebrand, C.; Laborde, D. *High Fertilizer Prices Contribute to Rising Global Food Security Concerns*; IFPRI—International Food Policy Research Institute: Washington, DC, USA, 2022.
8. Elias, E.H.; Flynn, R.; Idowu, O.J.; Reyes, J.; Sanogo, S.; Schutte, B.J.; Smith, R.; Steele, C.; Sutherland, C. Crop Vulnerability to Weather and Climate Risk: Analysis of Interacting Systems and Adaptation Efficacy for Sustainable Crop Production. *Sustainability* **2019**, *11*, 6619. [CrossRef]
9. Liu, Z.; Wang, S.; Zhang, Y.; Feng, Y.; Liu, J.; Zhu, H. Artificial Intelligence in Food Safety: A Decade Review and Bibliometric Analysis. *Foods* **2023**, *12*, 1242. [CrossRef] [PubMed]
10. Kutyauripo, I.; Rushambwa, M.; Chiwazi, L. Artificial intelligence applications in the agrifood sectors. *J. Agric. Food Res.* **2023**, *11*, 100502. [CrossRef]
11. Soori, M.; Arezoo, B.; Dastres, R. Artificial intelligence, machine learning and deep learning in advanced robotics, a review. *Cogn. Robot.* **2023**, *3*, 54–70. [CrossRef]
12. Sarker, I.H. AI-Based Modeling: Techniques, Applications and Research Issues Towards Automation, Intelligent and Smart Systems. *SN Comput. Sci.* **2022**, *3*, 158. [CrossRef]
13. Yokoi-Arai, M. *The Impact of Big Data and Artificial Intelligence (AI) in the Insurance Sector*; OECD: Paris, France, 2020.
14. Manning, L.; Brewer, S.; Craigon, P.J.; Frey, J.; Gutierrez, A.; Jacobs, N.; Kanza, S.; Munday, S.; Sacks, J.; Pearson, S. Artificial intelligence and ethics within the food sector: Developing a common language for technology adoption across the supply chain. *Trends Food Sci. Technol.* **2022**, *125*, 33–42. [CrossRef]
15. UNEP; FAO; UNDP. *Rethinking Our Food Systems: A Guide for Multi-Stakeholder Collaboration*; UNEP: Nairobi, Kenya; FAO: Rome, Italy; UNDP: New York, NY, USA, 2023.
16. Holzworth, D.P.; Snow, V.; Janssen, S.; Athanasiadis, I.N.; Donatelli, M.; Hoogenboom, G.; White, J.W.; Thorburn, P. Agricultural production systems modelling and software: Current status and future prospects. *Environ. Model. Softw.* **2015**, *72*, 276–286. [CrossRef]
17. Müller, B.; Hoffmann, F.; Heckelei, T.; Müller, C.; Hertel, T.W.; Polhill, J.G.; van Wijk, M.; Achterbosch, T.; Alexander, P.; Brown, C.; et al. Modelling food security: Bridging the gap between the micro and the macro scale. *Glob. Environ. Chang.* **2020**, *63*, 102085. [CrossRef]
18. Antle, J.M.; Jones, J.W.; Rosenzweig, C. Next generation agricultural system models and knowledge products: Synthesis and strategy. *Agric. Syst.* **2017**, *155*, 179–185. [CrossRef]

19. Verburg, P.H.; Alexander, P.; Evans, T.; Magliocca, N.R.; Malek, Z.; DA Rounsevell, M.; van Vliet, J. Beyond land cover change: Towards a new generation of land use models. *Curr. Opin. Environ. Sustain.* **2019**, *38*, 77–85. [CrossRef]
20. Kansou, K.; Laurier, W.; Charalambides, M.N.; Della-Valle, G.; Djekic, I.; Feyissa, A.H.; Marra, F.; Thomopoulos, R.; Bredeweg, B. Food modelling strategies and approaches for knowledge transfer. *Trends Food Sci. Technol.* **2022**, *120*, 363–373. [CrossRef]
21. van Wijk, M.T. From global economic modelling to household level analyses of food security and sustainability: How big is the gap and can we bridge it? *Food Policy* **2014**, *49*, 378–388. [CrossRef]
22. Lin, H.-Y. Large-Scale Artificial Intelligence Models. *Computer* **2022**, *55*, 76–80. [CrossRef]
23. Nasr-Azadani, E.; Wardrop, D.; Brooks, R. Is the rapid development of visualization techniques enhancing the quality of public participation in natural resource policy and management? A systematic review. *Landsc. Urban Plan.* **2022**, *228*, 104586. [CrossRef]
24. Mössinger, J.; Troost, C.; Berger, T. Bridging the gap between models and users: A lightweight mobile interface for optimized farming decisions in interactive modeling sessions. *Agric. Syst.* **2022**, *195*, 103315. [CrossRef]
25. Wens, M.L.K. *Farmers Facing Droughts: Capturing Adaptation Dynamics in Disaster Risk Models*; VRIJE Universiteit Amsterdam: Amsterdam, The Netherlands, 2022.
26. Bousquet, F.; Castella, J.; Christophe Trebuil, G.; Boissau, S.; Kam, S.P. The Use of Multi-agent Simulations in Companionable Modelling Approach for Agro-ecosystem Management. In Proceedings of the Contribution to the International Workshop "Integrated Management for Sustainable Agriculture, Forestry and Fisheries, Cali, Colombia, 28–31 August 2001; pp. 28–31.
27. Dafoe, A.; Bachrach, Y.; Hadfield, G.; Horvitz, E.; Larson, K.; Graepel, T. Cooperative AI: Machines must learn to find common ground. *Nature* **2021**, *593*, 33–36. [CrossRef]
28. Perrone, D.; Rohde, M.M.; Wagner, C.H.; Anderson, R.; Arthur, S.; Atume, N.; Brown, M.; Esaki-Kua, L.; Fernandez, M.G.; Garvey, K.A.; et al. Stakeholder integration predicts better outcomes from groundwater sustainability policy. *Nat. Commun.* **2023**, *14*, 3793. [CrossRef]
29. Halbe, J.; Holtz, G.; Ruutu, S. Participatory modeling for transition governance: Linking methods to process phases. *Environ. Innov. Soc. Transit.* **2020**, *35*, 60–76. [CrossRef]
30. van Wijk, M.; Rufino, M.; Enahoro, D.; Parsons, D.; Silvestri, S.; Valdivia, R.; Herrero, M. Farm household models to analyse food security in a changing climate: A review. *Glob. Food Secur.* **2014**, *3*, 77–84. [CrossRef]
31. Huber, R.; Bakker, M.; Balmann, A.; Berger, T.; Bithell, M.; Brown, C.; Grêt-Regamey, A.; Xiong, H.; Le, Q.B.; Mack, G.; et al. Representation of decision-making in European agricultural agent-based models. *Agric. Syst.* **2018**, *167*, 143–160. [CrossRef]
32. Reidsma, P.; Janssen, S.; Jansen, J.; van Ittersum, M.K. On the development and use of farm models for policy impact assessment in the European Union—A review. *Agric. Syst.* **2018**, *159*, 111–125. [CrossRef]
33. Mouratiadou, I.; Latka, C.; van der Hilst, F.; Müller, C.; Berges, R.; Bodirsky, B.L.; Ewert, F.; Faye, B.; Heckelei, T.; Hoffmann, M.; et al. Quantifying sustainable intensification of agriculture: The contribution of metrics and modelling. *Ecol. Indic.* **2021**, *129*, 107870. [CrossRef]
34. van der Linden, A.; de Olde, E.M.; Mostert, P.F.; de Boer, I.J. A review of European models to assess the sustainability performance of livestock production systems. *Agric. Syst.* **2020**, *182*, 102842. [CrossRef]
35. Laborde, D.; Majeed, F.; Tokgoz, S.; Torero, M. Long-Term Drivers of Food and Nutrition Security. IFPRI Discussion Paper Series. 2016. Available online: http://www.ifpri.org/publication/long-term-drivers-food-and-nutrition-security-0 (accessed on 26 September 2023).
36. Burchi, F.; De Muro, P. From food availability to nutritional capabilities: Advancing food security analysis. *Food Policy* **2016**, *60*, 10–19. [CrossRef]
37. Coates, J. Build it back better: Deconstructing food security for improved measurement and action. *Glob. Food Secur.* **2013**, *2*, 188–194. [CrossRef]
38. Herrero, M.; Thornton, P.K.; Power, B.; Bogard, J.R.; Remans, R.; Fritz, S.; Gerber, J.S.; Nelson, G.; See, L.; Waha, K.; et al. Farming and the geography of nutrient production for human use: A transdisciplinary analysis. *Lancet Planet. Health* **2017**, *1*, e33–e42. [CrossRef]
39. FAO. Food Security. Policy Brief June 2006, Issue 2. Policy Brief. 2006. Available online: https://reliefweb.int/report/world/policy-brief-food-security-issue-2-june-2006 (accessed on 7 March 2023).
40. Peng, W.; Berry, E.M. The concept of food security. *Encycl. Food Secur. Sustain.* **2018**, *2*, 1–7.
41. Headey, D.; Ecker, O. Rethinking the measurement of food security: From first principles to best practice. *Food Secur.* **2013**, *5*, 327–343. [CrossRef]
42. Stamoulis, K.; Zezza, A. *A Conceptual Framework for National Agricultural, Rural Development, and Food Security Strategies and Policies*; AgEcon Search: Washington, DC, USA, 2003.
43. Riahi, K.; Van Vuuren, D.P.; Kriegler, E.; Edmonds, J.; O'Neill, B.C.; Fujimori, S.; Bauer, N.; Calvin, K.; Dellink, R.; Fricko, O.; et al. The Shared Socioeconomic Pathways and their energy, land use, and greenhouse gas emissions implications: An overview. *Glob. Environ. Chang.* **2017**, *42*, 153–168. [CrossRef]
44. How, M.-L.; Chan, Y.J.; Cheah, S.-M. Predictive Insights for Improving the Resilience of Global Food Security Using Artificial Intelligence. *Sustainability* **2020**, *12*, 6272. [CrossRef]
45. Christin, S.; Hervet, É.; LeComte, N. Applications for deep learning in ecology. *Methods Ecol. Evol.* **2019**, *10*, 1632–1644. [CrossRef]
46. Voight, C.; Hernandez-Aguilar, K.; Garcia, C.; Gutierrez, S. Predictive Modeling of Future Forest Cover Change Patterns in Southern Belize. *Remote Sens.* **2019**, *11*, 823. [CrossRef]

47. Yang, L.; Driscol, J.; Sarigai, S.; Wu, Q.; Chen, H.; Lippitt, C.D. Google Earth Engine and Artificial Intelligence (AI): A Comprehensive Review. *Remote Sens.* **2022**, *14*, 3253. [CrossRef]
48. Schreinemachers, P.; Berger, T. Multi-Agent Systems for the Simulation of Land Use Change and Policy Interventions. Modelling and Simulation. 2008. Available online: https://citeseerx.ist.psu.edu/document?repid=rep1&type=pdf&doi=e4ddb5dd3890070a8209c61e4d9de2248c0c8f60 (accessed on 26 September 2023).
49. Salvini, G.; Ligtenberg, A.; van Paassen, A.; Bregt, A.; Avitabile, V.; Herold, M. REDD+ and climate smart agriculture in landscapes: A case study in Vietnam using companion modelling. *J. Environ. Manag.* **2016**, *172*, 58–70. [CrossRef]
50. Gonçalves, C.; Honrado, J.P.; Cerejeira, J.; Sousa, R.; Fernandes, P.M.; Vaz, A.S.; Alves, M.; Araújo, M.; Carvalho-Santos, C.; Fonseca, A.; et al. On the development of a regional climate change adaptation plan: Integrating model-assisted projections and stakeholders' perceptions. *Sci. Total Environ.* **2022**, *805*, 150320. [CrossRef]
51. Selbonne, S.; Guindé, L.; Belmadani, A.; Bonine, C.; Causeret, F.L.; Duval, M.; Sierra, J.; Blazy, J. Designing scenarios for upscaling climate-smart agriculture on a small tropical island. *Agric. Syst.* **2022**, *199*, 103408. [CrossRef]
52. Ermolieva, T.Y.; Ermoliev, Y.M.; Havlík, P.; Mosnier, A.; Leclere, D.; Fritz, S.; Valin, H.; Obersteiner, M.; Kyryzyuk, S.V.; Borodina, O.M. Dynamic Merge of the Global and Local Models for Sustainable Land Use Planning with Regard for Global Projections from GLOBIOM and Local Technical–Economic Feasibility and Resource Constraints. *Cybern. Syst. Anal.* **2017**, *53*, 176–185. [CrossRef]
53. Antle, J.M.; Stoorvogel, J.J.; Valdivia, R.O. New parsimonious simulation methods and tools to assess future food and environmental security of farm populations. *Philos. Trans. R. Soc. B Biol. Sci.* **2014**, *369*, 20120280. [CrossRef]
54. Gough, D.; Thomas, J.; Oliver, S. Clarifying differences between reviews within evidence ecosystems. *Syst. Rev.* **2019**, *8*, 170. [CrossRef]
55. Moher, D.; Liberati, A.; Tetzlaff, J.; Altman, D.G.; PRISMA Group. Preferred reporting items for systematic reviews and meta-analyses: The PRISMA statement. *PLoS Med.* **2009**, *6*, e1000097. [CrossRef] [PubMed]
56. Gebrehiwot, A.A.; Hashemi-Beni, L.; Kurkalova, L.A.; Liang, C.L.; Jha, M.K. Using ABM to Study the Potential of Land Use Change for Mitigation of Food Deserts. *Sustainability* **2022**, *14*, 9715. [CrossRef]
57. Feng, P. Modelling Impacts of Climate and Weather Extremes on Wheat Cropping Systems across New South Wales. Doctoral Dissertation, University of Technology Sydney, Sydney, Australia, 2020.
58. Gibson, M.; Pereira, J.P.; Slade, R.; Rogelj, J. Agent-Based Modelling of Future Dairy and Plant-Based Milk Consumption for UK Climate Targets. *J. Artif. Soc. Soc. Simul.* **2021**, *25*, 2. [CrossRef]
59. Hampf, A.C.; Carauta, M.; Latynskiy, E.; Libera, A.A.; Monteiro, L.; Sentelhas, P.; Troost, C.; Berger, T.; Nendel, C. The biophysical and socio-economic dimension of yield gaps in the southern Amazon—A bio-economic modelling approach. *Agric. Syst.* **2018**, *165*, 1–13. [CrossRef]
60. Carauta, M.; Parussis, J.; Hampf, A.; Libera, A.; Berger, T. No more double cropping in Mato Grosso, Brazil? Evaluating the potential impact of climate change on the profitability of farm systems. *Agric. Syst.* **2021**, *190*, 103104. [CrossRef]
61. Dobbie, S.; Schreckenberg, K.; Dyke, J.G.; Schaafsma, M.; Balbi, S. Agent-Based Modelling to Assess Community Food Security and Sustainable Livelihoods. *J. Artif. Soc. Soc. Simul.* **2018**, *21*, 9. [CrossRef]
62. Wossen, T.; Berger, T.; Haile, M.G.; Troost, C. Impacts of climate variability and food price volatility on household income and food security of farm households in East and West Africa. *Agric. Syst.* **2018**, *163*, 7–15. [CrossRef]
63. Balbi, S.; Alvarez-Rodriguez, U.; Latora, V.; Antonioni, A.; Villa, F. A game theory model to explore the role of cooperation and diversity in community food security: The case of Southern Malawi. *Reg. Environ. Chang.* **2020**, *20*, 63. [CrossRef]
64. Berre, D.; Diarisso, T.; Andrieu, N.; Le Page, C.; Corbeels, M. Biomass flows in an agro-pastoral village in West-Africa: Who benefits from crop residue mulching? *Agric Syst.* **2021**, *187*, 102981. [CrossRef]
65. Brinkmann, K.; Kübler, D.; Liehr, S.; Buerkert, A. Agent-based modelling of the social-ecological nature of poverty traps in southwestern Madagascar. *Agric. Syst.* **2021**, *190*, 103125. [CrossRef]
66. Jagustović, R.; Papachristos, G.; Zougmoré, R.B.; Kotir, J.H.; Kessler, A.; Ouédraogo, M.; Ritsema, C.J.; Dittmer, K.M. Better before worse trajectories in food systems? An investigation of synergies and trade-offs through climate-smart agriculture and system dynamics. *Agric. Syst.* **2021**, *190*, 103131. [CrossRef]
67. Alexander, S.; Block, P. Integration of seasonal precipitation forecast information into local-level agricultural decision-making using an agent-based model to support community adaptation. *Clim. Risk Manag.* **2022**, *36*, 100417. [CrossRef]
68. Bharwani, S. Adaptive Knowledge Dynamics and Emergent Artificial Societies: Ethnographically Based Multi-Agent Simulations of Behavioural Adaptation in Agro-Climatic Systems. Doctoral Thesis, University of Kent, Kent, UK, 2004.
69. Aurbacher, J.; Parker, P.S.; Sánchez, G.A.C.; Steinbach, J.; Reinmuth, E.; Ingwersen, J.; Dabbert, S. Influence of climate change on short term management of field crops—A modelling approach. *Agric. Syst.* **2013**, *119*, 44–57. [CrossRef]
70. Huber, R.; Rebecca, S.; François, M.; Hanna, B.S.; Dirk, S.; Robert, F. Interaction effects of targeted agri-environmental payments on non-marketed goods and services under climate change in a mountain region. *Land Use Policy* **2017**, *66*, 49–60. [CrossRef]
71. Grashof-Bokdam, C.J.; Cormont, A.; Polman, N.B.P.; Westerhof, E.J.G.M.; Franke, J.G.J.; Opdam, P.F.M. Modelling shifts between mono- and multifunctional farming systems: The importance of social and economic drivers. *Landsc. Ecol.* **2017**, *32*, 595–607. [CrossRef]
72. Aivazidou, E.; Tsolakis, N. Investigating dynamic interconnections between organic farming adoption and freshwater sustainability. *J. Environ. Manag.* **2021**, *294*, 112896. [CrossRef]

73. Saqalli, M.; Gérard, B.; Bielders, C.; Defourny, P. Targeting rural development interventions: Empirical agent-based modeling in Nigerien villages. *Agric. Syst.* **2011**, *104*, 354–364. [CrossRef]
74. Warner, K.; Afifi, T. Where the rain falls: Evidence from 8 countries on how vulnerable households use migration to manage the risk of rainfall variability and food insecurity. *Clim. Dev.* **2014**, *6*, 1–17. [CrossRef]
75. Fust, P.; Schlecht, E. Importance of timing: Vulnerability of semi-arid rangeland systems to increased variability in temporal distribution of rainfall events as predicted by future climate change. *Ecol. Model.* **2022**, *468*, 109961. [CrossRef]
76. Yuan, C.; Liu, L.; Qi, X.; Fu, Y.; Ye, J. Assessing the impacts of the changes in farming systems on food security and environmental sustainability of a Chinese rural region under different policy scenarios: An agent-based model. *Environ. Monit. Assess.* **2017**, *189*, 322. [CrossRef] [PubMed]
77. Badillo-Márquez, A.E.; Aguilar-Lasserre, A.A.; Miranda-Ackerman, M.A.; Sandoval-González, O.O.; Villanueva-Vásquez, D.; Posada-Gómez, R. An Agent-Based Model-Driven Decision Support System for Assessment of Agricultural Vulnerability of Sugarcane Facing Climatic Change. *Mathematics* **2021**, *9*, 3061. [CrossRef]
78. Grundy, M.J.; Bryan, B.A.; Nolan, M.; Battaglia, M.; Hatfield-Dodds, S.; Connor, J.D.; Keating, B.A. Scenarios for Australian agricultural production and land use to 2050. *Agric. Syst.* **2016**, *142*, 70–83. [CrossRef]
79. Chen, G.; Li, X.; Liu, X.; Chen, Y.; Liang, X.; Leng, J.; Xu, X.; Liao, W.; Qiu, Y.; Wu, Q.; et al. Global projections of future urban land expansion under shared socioeconomic pathways. *Nat. Commun.* **2020**, *11*, 537. [CrossRef]
80. Li, X.; Jousset, A.; de Boer, W.; Carrión, V.J.; Zhang, T.; Wang, X.; Kuramae, E.E. Legacy of land use history determines reprogramming of plant physiology by soil microbiome. *ISME J.* **2018**, *13*, 738–751. [CrossRef]
81. Feuerbacher, A.; McDonald, S.; Dukpa, C.; Grethe, H. Seasonal rural labor markets and their relevance to policy analyses in developing countries. *Food Policy* **2020**, *93*, 101875. [CrossRef]
82. Wossen, T.; Berger, T. Climate variability, food security and poverty: Agent-based assessment of policy options for farm households in Northern Ghana. *Environ. Sci. Policy* **2015**, *47*, 95–107. [CrossRef]
83. Alaliyat, S.; Yndestad, H.; Davidsen, P.I. An agent-based approach for predicting patterns of pathogen transmission between aquaculture sites in the Norwegian fjords. *Aquaculture* **2019**, *505*, 98–111. [CrossRef]
84. Williams, T.; Guikema, S.; Brown, D.; Agrawal, A. Resilience and equity: Quantifying the distributional effects of resilience-enhancing strategies in a smallholder agricultural system. *Agric. Syst.* **2020**, *182*, 102832. [CrossRef]
85. Saqalli, M.; Gérard, B.; Bielders, C.; Defourny, P. Testing the impact of social forces on the evolution of Sahelian farming systems: A combined agent-based modeling and anthropological approach. *Ecol. Model.* **2010**, *221*, 2714–2727. [CrossRef]
86. Ghaffari, A.; Bunch, M.J.; MacRae, R.J.; Zhao, S.J. Socio-economic Support Optimization for Transition from Conventional to Organic Farming Using a Spatiotemporal Agent-based Model. *Int. J. Interdiscip. Environ. Stud.* **2015**, *8*, 13–25. [CrossRef]
87. Groeneveld, A.; Bakker, M.; Peerlings, J.; Heijman, W. Complex dynamics in the uptake of new farming practices: A case study for organic waste application. *J. Environ. Plan. Manag.* **2019**, *62*, 818–842. [CrossRef]
88. Bazzana, D.; Foltz, J.; Zhang, Y. Impact of climate smart agriculture on food security: An agent-based analysis. *Food Policy* **2022**, *111*, 102304. [CrossRef]
89. Carauta, M.; Grovermann, C.; Heidenreich, A.; Berger, T. How eco-efficient are crop farms in the Southern Amazon region? Insights from combining agent-based simulations with robust order-m eco-efficiency estimation. *Sci. Total Environ.* **2022**, *819*, 153072. [CrossRef] [PubMed]
90. Wens, M.; Veldkamp, T.I.E.; Mwangi, M.; Johnson, J.M.; Lasage, R.; Haer, T.; Aerts, J.C.J.H. Simulating Small-Scale Agricultural Adaptation Decisions in Response to Drought Risk: An Empirical Agent-Based Model for Semi-Arid Kenya. *Front. Water* **2020**, *2*, 15. [CrossRef]
91. Grill, G.O. Spatial Agent-Based Modelling Applied to Irrigation Agriculture Dynamics in the Choapa-Valley, Chile. Master's Thesis, Paris Lodron University Salzburg, Nuremberg, Germany, 2005.
92. Jafary, F. Participatory Modelling Platform for Groundwater Irrigation Management with Local Farmers in Iran (Kashan). Doctoral Dissertation, University of Birmingham, Birmingham, UK, 2016.
93. Olayide, O.E.; Sangare, S.; Koo, J.; Xie, H. *Targeting Small-Scale Irrigation Investments Using Agent-Based Modeling: Case Studies in Mali and Niger*; Center for Development Research: Bonn, Germany, 2020. [CrossRef]
94. Akpoti, K.; Higginbottom, T.P.; Foster, T.; Adhikari, R.; Zwart, S.J. Mapping land suitability for informal, small-scale irrigation development using spatial modelling and machine learning in the Upper East Region, Ghana. *Sci. Total Environ.* **2022**, *803*, 149959. [CrossRef]
95. Priyadarshini, K.N.; Kumar, M.; Kumaraswamy, K. Identification of food insecure zones using remote sensing and artificial intelligence techniques. *ISPRS Int. Arch. Photogramm. Remote Sens. Spat. Inf. Sci.* **2018**, *XLII-5*, 659–664. [CrossRef]
96. Fujimori, S.; Hasegawa, T.; Rogelj, J.; Su, X.; Havlik, P.; Krey, V.; Takahashi, K.; Riahi, K. Inclusive climate change mitigation and food security policy under 1.5 °C climate goal. *Environ. Res. Lett.* **2018**, *13*, 074033. [CrossRef]
97. Peña-Lévano, L.M.; Taheripour, F.; Tyner, W.E. Climate Change Interactions with Agriculture, Forestry Sequestration, and Food Security. *Environ. Resour. Econ.* **2019**, *74*, 653–675. [CrossRef]
98. Nicholson, C.F.; Stephens, E.C.; Kopainsky, B.; Thornton, P.K.; Jones, A.D.; Parsons, D.; Garrett, J. Food security outcomes in agricultural systems models: Case examples and priority information needs. *Agric. Syst.* **2021**, *188*, 103030. [CrossRef]
99. Wens, M.L.K.; Van Loon, A.F.; Veldkamp, T.I.E.; Aerts, J.C.J.H. Education, financial aid, and awareness can reduce smallholder farmers' vulnerability to drought under climate change. *Nat. Hazards Earth Syst. Sci.* **2022**, *22*, 1201–1232. [CrossRef]

100. Mejía, G.; García-Díaz, C. Market-level effects of firm-level adaptation and intermediation in networked markets of fresh foods: A case study in Colombia. *Agric Syst.* **2018**, *160*, 132–142. [CrossRef]
101. Dijkxhoorn, Y.; Plaisier, C.; Verwaart, T.; Van Wagenberg, C.; Ruben, R. Trusted sorghum: Simulating interactions in the sorghum value chain in Kenya using games and agent-based modelling. *J. Dev. Eff.* **2019**, *11*, 146–164. [CrossRef]
102. Nattassha, R.; Simatupang, T.M.; Handayati, Y. *Design of Circular Supply Chain Agent-Based Model: Case of Organic Fertilizer Producer*; Atlantis Press: Amsterdam, The Netherlands, 2019; pp. 347–352.
103. Nguyen, H.K.; Chiong, R.; Chica, M.; Middleton, R.; Pham, D.T.K. Contract Farming in the Mekong Delta's Rice Supply Chain: Insights from an Agent-Based Modeling Study. *J. Artif. Soc. Soc. Simul.* **2019**, *22*, 1. [CrossRef]
104. Namany, S.; Govindan, R.; Alfagih, L.; McKay, G.; Al-Ansari, T. Sustainable food security decision-making: An agent-based modelling approach. *J. Clean. Prod.* **2020**, *255*, 120296. [CrossRef]
105. Lu, L.; Nguyen, R.; Rahman, M.M.; Winfree, J. *Demand Shocks and Supply Chain Resilience: An Agent Based Modelling Approach and Application to the Potato Supply Chain*; National Bureau of Economic Research: Cambridge, MA, USA, 2021.
106. Williams, T.G.; Brown, D.G.; Agrawal, A.; Guikema, S.D. Let the farmer decide: Examining smallholder autonomy in large-scale land acquisitions with an agent-based model. *Environ. Res. Lett.* **2021**, *16*, 105011. [CrossRef]
107. Bazzana, D.; Zaitchik, B.; Gilioli, G. Impact of water and energy infrastructure on local well-being: An agent-based analysis of the water-energy-food nexus. *Struct. Chang. Econ. Dyn.* **2020**, *55*, 165–176. [CrossRef]
108. Haji, M.; Govindan, R.; Al-Ansari, T. A computational modelling approach based on the 'Energy–Water–Food nexus node' to support decision-making for sustainable and resilient food security. *Comput. Chem. Eng.* **2022**, *163*, 107846. [CrossRef]
109. Tian, Q.; Holland, J.H.; Brown, D.G. Social and economic impacts of subsidy policies on rural development in the Poyang Lake Region, China: Insights from an agent-based model. *Agric. Syst.* **2016**, *148*, 12–27. [CrossRef]
110. Dou, Y. The Impacts of Cash Transfer Programs on Rural Livelihoods: A Study of Caboclos in the Brazilian Amazon Estuary Region. Doctoral Thesis, University of Waterloo, Waterloo, ON, Canada, 2016.
111. Reilly, J.A.; Dawson, T.P.; Matthews, R.B.; Smith, P.; Musk, C.C.; Potts, J.M.; Polhill, J.G. Projecting the effect of crop yield increases, dietary change and different price scenarios on land use under two different state security regimes. *Int. J. Agric. Sustain.* **2021**, *19*, 288–304. [CrossRef]
112. Natalini, D.; Bravo, G.; Jones, A.W. Global food security and food riots—An agent-based modelling approach. *Food Secur.* **2017**, *11*, 1153–1173. [CrossRef]
113. *Use of Agent-Based Models in Characterizing Farm Types and Evolvement in Smallholder Dairy Systems*; Nelson Mandela African Institution of Science and Technology: Arusha, Tanzania, 2020. Available online: https://dspace.nm-aist.ac.tz/handle/20.500.12479/895 (accessed on 26 September 2023).
114. Thomopoulos, R.; Salliou, N.; Abreu, C.; Cohen, V.; Fouqueray, T. Re-duced Meat Consumption: From Multicriteria Argument Modelling to Agent-Based Social Simulation. *Int. J. Food Stud.* **2021**, *2021*, 133–149. [CrossRef]
115. Lloyd, S.J.; Chalabi, Z. Climate change, hunger and rural health through the lens of farming styles: An agent-based model to assess the potential role of peasant farming. *PLoS ONE* **2021**, *16*, e0246788. [CrossRef] [PubMed]
116. Hailegiorgis, A.; Crooks, A.; Cioffi-Revilla, C. An Agent-Based Model of Rural Households' Adaptation to Climate Change. *J. Artif. Soc. Soc. Simul.* **2018**, *21*, 4. [CrossRef]
117. Koh, K.; Reno, R.; Hyder, A. Examining disparities in food accessibility among households in Columbus, Ohio: An agent-based model. *Food Secur.* **2019**, *11*, 317–331. [CrossRef]
118. Li, J.; Oyana, T.J.; Mukwaya, P.I. An examination of historical and future land use changes in Uganda using change detection methods and agent-based modelling. *Afr. Geogr. Rev.* **2016**, *35*, 247–271. [CrossRef]
119. Pant, H.; Tulpulé, V.; Fisher, B.S. The Global Trade and Environment Model: A Projection of Non-Steady State Data Using Intertemporal GTEM. In Proceedings of the Fifth Annual Conference on Global Economic Analysis, Taipei, Taiwan, 5–7 June 2002.
120. Hayward, J.A.; Graham, P.W. A global and local endogenous experience curve model for projecting future uptake and cost of electricity generation technologies. *Energy Econ.* **2013**, *40*, 537–548. [CrossRef]
121. Nainggolan, D.; Termansen, M.; Fleskens, L.; Hubacek, K.; Reed, M.S.; de Vente, J.; Boix-Fayos, C. What does the future hold for semi-arid Mediterranean agro-ecosystems?—Exploring cellular automata and agent-based trajectories of future land-use change. *Appl. Geogr.* **2012**, *35*, 474–490. [CrossRef]
122. Gittins, P.; McElwee, G.; Tipi, N. Discrete event simulation in livestock management. *J. Rural. Stud.* **2020**, *78*, 387–398. [CrossRef]
123. Guo, M.; van Dam, K.H.; Touhami, N.O.; Nguyen, R.; Delval, F.; Jamieson, C.; Shah, N. Multi-level system modelling of the resource-food-bioenergy nexus in the global south. *Energy* **2020**, *197*, 117196. [CrossRef]
124. Barba-Escoto, L.; van Wijk, M.T.; López-Ridaura, S. Non-linear Interactions Driving Food Security of Smallholder Farm Households in the Western Highlands of Guatemala. *Front. Sustain. Food Syst.* **2020**, *4*, 51. [CrossRef]
125. Kim, N.; Na, S.-I.; Park, C.-W.; Huh, M.; Oh, J.; Ha, K.-J.; Cho, J.; Lee, Y.-W. An Artificial Intelligence Approach to Prediction of Corn Yields under Extreme Weather Conditions Using Satellite and Meteorological Data. *Appl. Sci.* **2020**, *10*, 3785. [CrossRef]
126. Wang, X.; Huang, J.; Feng, Q.; Yin, D. Winter Wheat Yield Prediction at County Level and Uncertainty Analysis in Main Wheat-Producing Regions of China with Deep Learning Approaches. *Remote Sens.* **2020**, *12*, 1744. [CrossRef]
127. Yang, R.; Ahmed, Z.U.; Schulthess, U.C.; Kamal, M.; Rai, R. Detecting functional field units from satellite images in smallholder farming systems using a deep learning based computer vision approach: A case study from Bangladesh. *Remote Sens. Appl. Soc. Environ.* **2020**, *20*, 100413. [CrossRef] [PubMed]

128. Afzaal, H.; Farooque, A.A.; Schumann, A.W.; Hussain, N.; McKenzie-Gopsill, A.; Esau, T.; Abbas, F.; Acharya, B. Detection of a Potato Disease (Early Blight) Using Artificial Intelligence. *Remote Sens.* **2021**, *13*, 411. [CrossRef]
129. Meng, L.; Liu, H.; Ustin, S.L.; Zhang, X. Predicting Maize Yield at the Plot Scale of Different Fertilizer Systems by Multi-Source Data and Machine Learning Methods. *Remote Sens.* **2021**, *13*, 3760. [CrossRef]
130. Al-Adhaileh, M.H.; Aldhyani, T.H. Artificial intelligence framework for modeling and predicting crop yield to enhance food security in Saudi Arabia. *PeerJ Comput. Sci.* **2022**, *8*, e1104. [CrossRef]
131. Belem, M.; Bazile, D.; Coulibaly, H. Simulating the Impacts of Climate Variability and Change on Crop Varietal Diversity in Mali (West-Africa) Using Agent-Based Modeling Approach. *J. Artif. Soc. Soc. Simul.* **2018**, *21*, e8. [CrossRef]
132. Ge, J.; Scalco, A.; Craig, T. Social Influence and Meat-Eating Behaviour. *Sustainability* **2022**, *14*, 7935. [CrossRef]
133. Bharwani, S.; Bithell, M.; Downing, T.E.; New, M.; Washington, R.; Ziervogel, G. Multi-agent modelling of climate outlooks and food security on a community garden scheme in Limpopo, South Africa. *Philos. Trans. R. Soc. B Biol. Sci.* **2005**, *360*, 2183–2194. [CrossRef]
134. Breisinger, C.; Ecker, O.; Al-Riffai, P.; Robertson, R.; Thiele, R.; Wiebelt, M. *Climate Change, Agricultural Production and Food Security: Evidence from Yemen*; Kiel Institute for the World Economy (IfW): Kiel, Germany, 2011; pp. 1–34.
135. Fincham, A. Modelling Resilience to Food Insecurity Malawi Case Study. Doctoral Dissertation, Stellenbosch University, Stellenbosch, South Africa, 2020.
136. Vanany, I.; Hajar, G.; Utami, N.M.C.; Jaelani, L.M. Modelling food security for staple protein in Indonesia using system dynamics approach. *Cogent Eng.* **2021**, *8*, 2003945. [CrossRef]
137. Musayev, S.; Mellor, J.; Walsh, T.; Anagnostou, E. Application of Agent-Based Modeling in Agricultural Productivity in Rural Area of Bahir Dar, Ethiopia. *Forecasting* **2022**, *4*, 349–370. [CrossRef]
138. Nöldeke, B.; Winter, E.; Laumonier, Y.; Simamora, T. Simulating Agroforestry Adoption in Rural Indonesia: The Potential of Trees on Farms for Livelihoods and Environment. *Land* **2021**, *10*, 385. [CrossRef]
139. Ambrosius, F.H.W. Transition through Markets: Understanding Diffusion of Added-Value Markets among Dutch Pig Farmers. Doctoral Dissertation, Wageningen University and Research, Wageningen, The Netherlands, 2021.
140. Sanga, U.; Park, H.; Wagner, C.H.; Shah, S.H.; Ligmann-Zielinska, A. How do farmers adapt to agricultural risks in northern India? An agent-based exploration of alternate theories of decision-making. *J. Environ. Manag.* **2021**, *298*, 113353. [CrossRef]
141. Dangles, O.; Rebaudo, F. Adaptive management in crop pest control in the face of climate variability: An agent-based modeling approach. *Ecol. Soc.* **2015**, *20*, 2. [CrossRef]
142. Paulausky, P. Application of an Agent-Based Modeling Approach to Study the Postharvest Loss of Food Grains in India. Master's Thesis, University of Illinois at Urbana-Champaign, Champaign, IL, USA, 2018.
143. Assfaw, T.W. Dissertation Climate Variability, Social Capital and Food Security in Sub-Saharan Africa: Household Level Assessment of Potential Impacts and Adaptation Options. Doctoral Dissertation, Universität Hohenheim, Hohenheim, Germany, 2015.
144. Bousquet, F.; Castella, J.C.; Trébuil, G.; Barnaud, C.; Boissau, S.; Suan, P.K. Using multi-agent systems in a companion modelling approach for agro-ecosystem management in South-east Asia. *Outlook Agric.* **2007**, *36*, 57–62. [CrossRef]
145. Grillot, M.; Vayssières, J.; Masse, D. Agent-based modelling as a time machine to assess nutrient cycling reorganization during past agrarian transitions in West Africa. *Agric. Syst.* **2018**, *164*, 133–151. [CrossRef]
146. Belem, M.; Saqalli, M. Development of an integrated generic model for multi-scale assessment of the impacts of agro-ecosystems on major ecosystem services in West Africa. *J. Environ. Manag.* **2017**, *202*, 117–125. [CrossRef]
147. Schreinemachers, P.; Berger, T.; Aune, J.B. Simulating soil fertility and poverty dynamics in Uganda: A bio-economic multi-agent systems approach. *Ecol. Econ.* **2007**, *64*, 387–401. [CrossRef]
148. Potchanasin, C. Simulation of the Sustainability of Farming Systems in Northern Thailand. 2008. Available online: http://opus.uni-hohenheim.de/volltexte/2008/306/ (accessed on 26 September 2023).
149. Lindner, A.; Jost, F.; Merino, M.V.; Reategui, N.; Pretzsch, J. Aligning Socio-economic Field Laboratories and Agent Based Models assessing local climate change adaptation measures of Andean farmers. *J. For. Landsc. Res.* **2017**, *2*, 7–17. [CrossRef]
150. Arnold, T. Linking Farm Economics and Hydrology: Model Integration for Watershed-Level Irrigation Management Applied to Chile Dissertation. 2009. Available online: http://opus.uni-hohenheim.de/volltexte/2010/499/ (accessed on 26 September 2023).
151. Reinmuth, E. Land Use Management under Climate Change: A Microeconomic Analysis with Emphasis on Risk. 2018. Available online: http://opus.uni-hohenheim.de/volltexte/2018/1473/ (accessed on 26 September 2023).
152. Fernandez-Mena, H.; MacDonald, G.K.; Pellerin, S.; Nesme, T. Co-benefits and Trade-Offs From Agro-Food System Redesign for Circularity: A Case Study With the FAN Agent-Based Model. *Front. Sustain. Food Syst.* **2020**, *4*, 41. [CrossRef]
153. Connor, J.D.; Bryan, B.A.; Nolan, M.; Stock, F.; Gao, L.; Dunstall, S.; Graham, P.; Ernst, A.; Newth, D.; Grundy, M.; et al. Modelling Australian land use competition and ecosystem services with food price feedbacks at high spatial resolution. *Environ. Model. Softw.* **2015**, *69*, 141–154. [CrossRef]
154. Mantzaris, A.V.; Rein, S.R.; Hopkins, A.D. Examining Collusion and Voting Biases Between Countries During the Eurovision Song Contest Since 1957. *arXiv* **2017**, arXiv:1705.06721. [CrossRef]
155. Kashka, F.M.; Sarvestani, Z.T.; Pirdashti, H.; Motevali, A.; Nadi, M. Predicting of Agro-Environmental Footprint with Artiicial Intelligence (Soybean Cultivation in Various Scenarios). 2022. Available online: https://www.researchsquare.com/article/rs-1098555/v1 (accessed on 26 September 2023).

156. Fernández, F.J.; Blanco, M.; Ponce, R.D.; Vásquez-Lavín, F.; Roco, L. Implications of climate change for semi-arid dualistic agriculture: A case study in Central Chile. *Reg. Environ. Chang.* **2018**, *19*, 89–100. [CrossRef]
157. Roco, L.; Poblete, D.; Meza, F.; Kerrigan, G. Farmers' Options to Address Water Scarcity in a Changing Climate: Case Studies from two Basins in Mediterranean Chile. *Environ. Manag.* **2016**, *58*, 958–971. [CrossRef]

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.

agriculture

MDPI

Review

Key Technologies of Plug Tray Seedling Transplanters in Protected Agriculture: A Review

Wei Liu [1,2], Shijie Tian [1], Qingyu Wang [1,2] and Huanyu Jiang [1,2,*]

[1] College of Biosystems Engineering and Food Science, Zhejiang University, Hangzhou 310058, China; 12013057@zju.edu.cn (W.L.); shijie_tian@zju.edu.cn (S.T.); 12013027@zju.edu.cn (Q.W.)
[2] Key Laboratory of Intelligent Equipment and Robotics for Agriculture, Hangzhou 310058, China
* Correspondence: hyjiang@zju.edu.cn; Tel.: +86-0571-88982170

Abstract: The process of plug tray seedling transplanting is a crucial step in protected agriculture production. Due to issues such as high labor intensity, poor consistency of work quality, and low efficiency, the application of automated transplanting machines has provided a solution to these issues. For the diversity of transplanting operations, various mechanical structures and technological applications have been developed for automated transplanting equipment. Therefore, this paper provides systematic research of current studies on the key transplanter technologies. Firstly, through an analysis of the types of transplanting operations, the technical requirements of automated transplanting equipment for different operation types are elucidated. Subsequently, the key technologies applied in transplanting machines are discussed from the perspectives of substrate physical characteristics, end effectors, integration of multiple end effectors, vision systems, and transplanting path planning. Moreover, an analysis is conducted on the advantages, disadvantages, and application scenarios of different research methods for each key technology. Lastly, the existing problems and technical difficulties of the transplanting machine are summarized, and future research directions are discussed. This analysis provides a valuable reference for further research and development in the field of transplanting machines for plug tray seedlings.

Keywords: transplanter; plug tray seedlings; protected agriculture; end effector; vision system

Citation: Liu, W.; Tian, S.; Wang, Q.; Jiang, H. Key Technologies of Plug Tray Seedling Transplanters in Protected Agriculture: A Review. *Agriculture* **2023**, *13*, 1488. https://doi.org/10.3390/agriculture13081488

Academic Editors: Jin Yuan and Zichen Huang

Received: 4 July 2023
Revised: 17 July 2023
Accepted: 25 July 2023
Published: 27 July 2023

Copyright: © 2023 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (https://creativecommons.org/licenses/by/4.0/).

1. Introduction

Vegetables are an indispensable component of the human diet, as they provide essential nutrition, such as fiber, vitamins, antioxidants, and minerals [1]. The increasing global population has resulted in a surge in demand for vegetables [2]. Protected agriculture, including indoor, greenhouse, and artificial light plant factory cultivation [3–5], offers a viable solution to meet this demand [6]. Protected agriculture employs techniques that control climatic factors, such as light, carbon dioxide content, temperature, and humidity, to create favorable conditions for plants [7,8]. This method provides high yield, good quality, continuous production, efficient resource utilization, and minimal impact on environmental changes [9–11], making it an excellent option for vegetable crop production [12].

Plug tray seedling technology is widely utilized in protected agriculture due to its high germination rate, neat growth, seed-saving benefits, and suitability for mechanized operation [13]. After a period of growth in a plug tray, the seedlings must be picked, replanted, transplanted, and sorted to foster optimal growth conditions [14]. Plug tray seedling transplanters are machines used for transplanting seedlings from the high-density plug tray to the low-density plug tray or flower pot. A complete set of transplanting equipment typically requires functions such as inferior seedling position recognition, tray conveying, tray hole positioning, substrate gripping, and seedling planting. Additionally, since the target of transplanting machine operations is the substrate with seedlings, it is necessary to ensure that the transplanting gripper does not cause any damage to the

seedlings during the transplanting process. Therefore, the design of transplanting machines is more complex compared to commonly used industrial automation equipment.

The research on transplanting equipment for vegetable production in protected agriculture started earlier in the 1980s and 1990s. In America, L.J. Kutz et al. [15] designed a nursery plant transplanter based on Puma560 industrial robot; K.C.TING et al. [16,17] developed a transplanter based on the SCARA robot. In the Czech Republic, HůLa et al. [18] conducted a comparative study on the transplanting effects of three different types of ABB robots. These early studies verified the feasibility of using industrial robots for tray seedling transplanting operations, but they did not conduct practical application research. In Korea, K.H.Ryu et al. [19] developed a transplanter with a CCD vision system, and Kang et al. [20] developed a transplanter with multiple claws based on the Cartesian coordinate manipulator. These transplanting machines have greatly improved the transplanting efficiency and success rate, and they basically met the requirements of practical applications. In Japan, Onosaka et al. [21] invented a transplanter composed of devices for transporting, grasping, selecting, planting, and seedling tray cleaning. Researchers in Japan have conducted extensive research on transplanting machines for field use [22,23] and have also developed transplanting machines for hydroponic cultivation [24,25], but there have been few studies on transplanting machines for plug tray seedlings in protected agriculture. After years of development, Europe has seen the emergence of mature commercial transplanting machines, especially in The Netherlands. The Dutch company Viscon [26] has extensive experience in facility horticulture and has produced a series of transplanters, such as Pic-O-Mat Blueline equipped with multiple patented end effectors, and the transplanting efficiency could reach 35,000 plants per hour. Viscon also has advanced vision systems, whereby crops can be effectively and reliably graded by detecting the volume, height, color, flower quantity, and other characteristics of a plant using a high-resolution camera. The transplanter produced by the TTA Company [27] in The Netherlands is widely used for transplanting vegetables, flowers, and other crops in facility horticulture, such as the PackPlanter Wireless series transplanter for transplanting, the FlexSorter series transplanters for grading, and the MidiPot and MidiCurve for potted flower transplanting. Other facility agricultural suppliers in The Netherlands, such as the Flier Systems Company [28], the ISO Group Company [29], the CODEMA Company [30], have also developed efficient transplanting equipment, which is widely used in practical production. Other companies, such as Urbanati [31] and TEA Project srl [32] in Italy, AgriNomix [33] and Bouldin & Lawson [34] in America, and Transplant System [35] in New Zealand, have all developed commercial transplanters for plug tray seedlings in protected agriculture. Although these commercial transplanting machines have matured in terms of functionality and design, they are designed for planting patterns and standards in a large-scale protected agriculture environment, especially in developed countries. However, there are differences in the production scale, planting mode, planting density, hardware configuration, and other aspects of protected agriculture in different countries. Transplanters have limitations in intelligence, applicability, and universality [36]. In addition, these complete sets of transplanting machines also face problems, such as high prices, high maintenance costs, and long maintenance cycles, which are still difficult for small and medium-sized facility vegetable production enterprises to accept. Therefore, Chinese researchers have developed many plug tray seedling transplanters suitable for their own planting patterns [13,37–41]. In addition to China, research on transplanting machines for potted seedlings in other developing countries has mainly focused on the application of transplanting machines in the field [42–46].

It is evident that the current commercial transplanting machines lack suitable applicability within the facility agricultural environment of developing countries. Transplanters incorporate multiple technologies, including mechanical, electrical, electronic, computer, machine vision, and intelligent sensing. Precisely due to the fact that transplanting machines are electromechanical integration devices encompassing multiple disciplinary technologies, there is a pressing need to scrutinize the pivotal technologies employed in existing

transplanting machines, enhance their structural integrity, optimize their functionalities, and mitigate costs. These endeavors aim to render the machines adaptable to a broader spectrum of planting conditions, thereby facilitating their widespread implementation. Several scholars have reviewed some aspects of the key technologies for transplanting machines, but their research content is not comprehensive enough. The differences between the recently published reviews of technologies in plug tray seedling transplanters and this paper were compared and shown in Table 1.

Table 1. Comparison of the current review paper with previous reviews.

References	Content								
	Transplanting Types	Research Status	Substrate Physical Properties	End Effector	Multiple End Effector Integration	Overall Structure	Vision System	Path Planning	Research Direction
[47]		√		√			√	√	√
[48]				√					√
[36]	√	√		√	√		√	√	√
[49]				√					√
[50]				√	√				√
ours	√	√	√	√	√	√	√	√	√

The rest of this paper is organized as follows. In Section 2, we analyze the different types of transplanting operations. In the Section 3, we summarize the key technologies in the design of plug tray seedling transplanters. Finally, we discuss the current problems in the design of plug tray seedling transplanters. The paper is concluded with a conclusion and the future research direction for plug tray seedling transplanters.

2. Overview of Transplanting Operation

Automated transplanting operations can be categorized into several types based on their specific seedling purposes. These include widening-spacing transplanting, replacing bad seedling transplanting, grading transplanting, and tray-to-pot transplanting. As a result, the structural principles of the transplanter vary depending on the type of transplanting operation. In the following paragraphs, we will provide an overview of each type of transplanting operation.

2.1. Types of Transplanting Operation

To maximize planting space, seedlings are typically planted in high-density plug trays with small hole spacing. However, as the plants grow, the crowded environment becomes unsuitable for further plant development. To address this issue, seedlings must be transplanted into larger trays with wider hole spacing, which improves their growth space, light utilization, and ventilation. This process is known as widening-spacing transplanting, as shown in Figure 1a. Several transplanters are available for this kind of operation, including the Pack Planter series from the TTA company in The Netherlands and the RW series from the Urbinati company in Italy.

Substandard seed quality, inaccurate seeding in plug trays, suboptimal temperature and humidity conditions, inadequate nutrition, as well as plant diseases and pests, collectively contribute to a suboptimal germination rate (ranging from 80% to 95%) for seedlings cultivated in plug trays [51]. This results in poor-quality seedlings or empty holes. To mitigate the risks associated with disease and pest outbreaks and enhance the utilization rate of plug trays, it is imperative to implement procedures for the identification and removal of unhealthy seedlings or empty holes, followed by the subsequent replanting with vigorous and disease-free seedlings. This type of operation is referred to as replacing bad seedling transplanting, as shown in Figure 1b. The Fix-O-Mat TIFS-IV from the Viscon company and the CombiFix series from the TTA company are both examples of this type of transplanter.

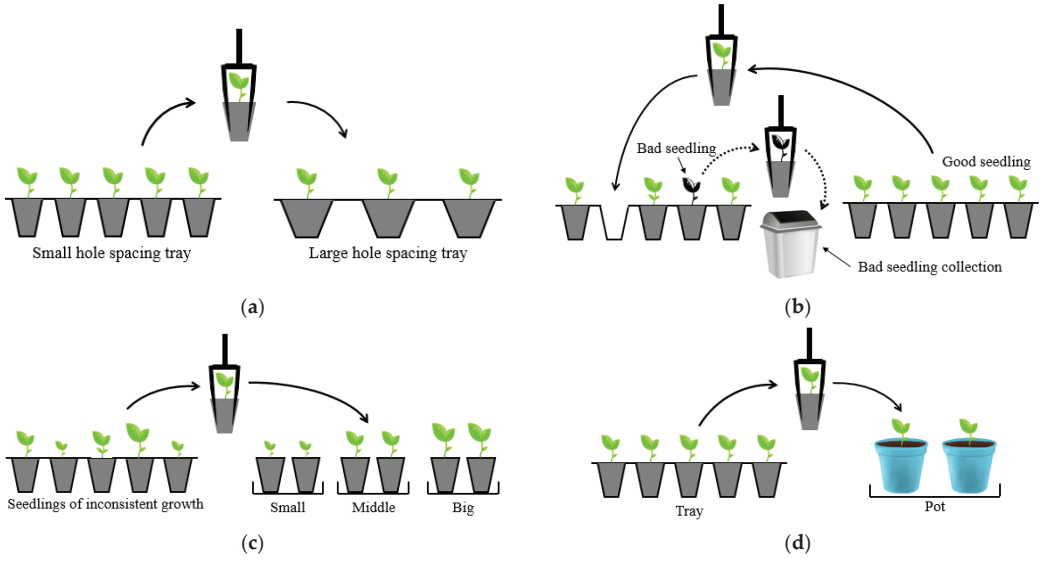


Figure 1. Types of transplanting: (**a**) widening-spacing transplanting, (**b**) replacement transplanting, (**c**) grading transplanting, (**d**) tray-to-pot transplanting.

The process of grading transplanting, as shown in Figure 1c, involves categorizing and grouping seedlings based on factors such as leaf size, leaf count, or plant height. This is typically performed to ensure that seedlings with similar characteristics are transplanted together, improving their uniformity and making them easier to manage. Grading can also be performed for marketing purposes, as seedlings with similar characteristics are often more attractive to buyers. The Select-O-Mat Phoenix from Viscon and the FlexSorter series from TTA are examples of grading transplanters.

Tray-to-pot transplanting is usually used for flower transplanting, where one or several flower seedlings from a plug tray are transplanted into a flower pot, as shown in Figure 1d. The Pic-O-Mat series from Viscon and the MidiCurve from TTA are examples of this type.

Some transplanting operations require a combination of the techniques mentioned above to meet specific process requirements. For example, TTA's FlexPlanter series can perform both widening-spacing and replacement transplanting, while Young Plant Sorter from Fliersystems can accomplish both grading and widening-spacing transplanting. The SPH-Sorter from Fliersystems can conduct replacement, widening-spacing, and grading transplanting.

2.2. *Technical Requirements for Transplanting Operations*

The structure and control methods of transplanters are determined by the types of transplanting operations performed. Therefore, the technical requirements of different types of transplanters vary accordingly.

The objective of widening-spacing transplanting is to augment the inter-plant spacing, necessitating the utilization of transplanter equipment equipped with multiple end effectors capable of adjusting the spacing between seedlings. When all seedlings in a tray are of good quality, the multiple end effectors can be lifted and gripped simultaneously, making the structural design and control methods of the end effectors and transplanting manipulators relatively simple. However, when the tray contains poor-quality seedlings or empty holes, the end effectors must independently lift and switch the grippers, making the overall structure and control of the transplanter more complex.

For replacement transplanting, the positions of poor-quality seedlings or empty holes in the tray are unknown, and thus, the end effectors must have the ability to stretch and

grip independently. Modern commercial transplanting machines usually use a parallel structure with multiple end effectors to increase work efficiency.

Grading transplanting requires a more precise visual system and an end effector capable of independent driving. The tray-to-pot transplanting principle is similar to widening-spacing transplanting, and the technical requirements are also similar. Table 2 summarizes the technical requirements of transplanters for different types of transplanting operations.

Table 2. Technical requirements of transplanters for different operation types.

Types	Change Spacing for End Effectors	Independent Control for End Effectors	Visual System
Widening-spacing transplanting	Yes	Needed or not	No
Replacement transplanting	Needed or not	Yes	Yes
Grading transplanting	Needed or not	Yes	Yes
Tray-to-pot transplanting	Yes	Needed or not	No

3. Key Technologies of Transplanting Machines

A plug tray seedling transplanter system consists of various components, including the end effectors, transplanting manipulator, conveying and positioning mechanism, visual system, and control system. The research of a tray seedling transplanting machine involves several stages. The first stage involves studying the physical properties of the seedling substrate to design and test the end effector. Next, the transplanting manipulator is designed according to the specific agronomic requirements. Finally, the entire machine is optimized to achieve all the necessary functions required for the transplanting operation. The system composition diagram of plug tray seedling transplanting machines for protected agriculture is depicted in Figure 2.

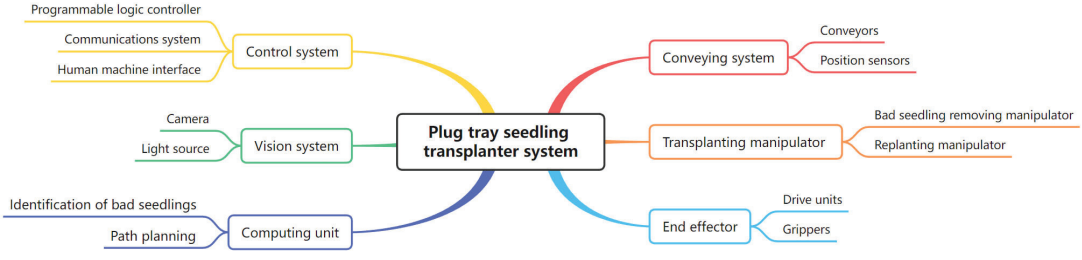


Figure 2. System composition diagram of plug tray seedling transplanter.

3.1. Physical Properties of Substrate

The physical properties of the plug tray seedling substrate are crucial in determining the success rate of the end effector's grasp. To optimize the design of the end effector, researchers have investigated this issue using methods such as force measurement platforms, simulation technology, and image technology. Their findings provide a theoretical foundation for the development of end effectors. A comparison of research methods for the physical properties of the substrate is shown in Table 3.

Table 3. Comparison of physical characteristics research methods.

Methods	Instruments/Tools	Purpose	Characteristics	Reference
Force measuring platform	Dynamometer	Measure the pulling force, penetration resistance, and clamping force of the gripper.	Cheap measurement platform needs to be designed, and the measurement data need to be recorded manually.	[52–54]
	Universal testing machine	Measure the tensile and compressive strength of the substrate, the pulling force, penetration resistance.	Simple operation, automatic recording of test data, moderate cost.	[55–57]
	Texture analyzer	Measure the matrix compression and creep test, the pulling force, penetration resistance.	Simple operation, high accuracy, automatic recording of test data, high cost.	[58–61]
	Direct shear apparatus	Measure the shear strength and shear stress of the substrate.	The test operation is complex, the test data are automatically recorded, high cost.	[62]
Simulation technology	FEM	Analyze the stress of the substrate and the gripper, simulate the damage of the substrate under different physical characteristic parameters.	Suitable for analyzing a substrate, which is easy to lose.	[63]
	DEM	Analyze the scattering of the substrate during the grasping process under different physical characteristic parameters.	Suitable for loose substrate.	[14,57]
Image technology	CT	Study the relationship between root density distribution, root micro-displacement, and the substrate crack expansion.	The equipment parameters need to be adjusted to obtain the three-dimensional image of root distribution.	[38,64,65]
	SEM	Study the substrate damage mechanism at the microscopic level of the internal composition and structure of the substrate.	Obtaining the internal section image of the substrate.	[66]

3.1.1. Force Measuring Platform

The force measuring platform can directly measure the force when the gripper acts on the substrate. Various instruments, such as force gauges, universal testing machines, texture analyzers, and direct shear instruments, are used to measure the tensile, compressive, or shear strength of substrates or seedlings, providing important data for optimizing the end effectors.

The dynamometer is commonly used for measuring various mechanical properties, including the penetration resistance of soil [53], the clamping force of the gripper [52], and the pulling force of the seedling [54]. The universal testing machine is often utilized for determining the substrate cohesion and tray adhesion [57], compression [55], tension [56], bending, and puncture. The texture analyzer is used to measure the compression [58–60], tearing, and surface puncture tests [61] of the samples. Additionally, the direct shear apparatus, is employed for measuring the shear strength of the substrate. By measuring the shear stress and shear displacement of the substrate, the cohesive force and internal friction angle of the substrate can be calculated [62].

3.1.2. Simulation Technology

Simulation technology can simulate the grasping process of seeding. By utilizing the finite element method (FEM) or discrete element method (DEM) simulation, researchers can simulate the grabbing process of substrate blocks by grippers, allowing for the calculation of the forces, deformation, and movement of both the gripper and substrate.

Several software, which can perform finite element analysis, such as Solidworks [63], Abaqus and Ansys [67]. These software can simulate the mechanical response of the needle force on the seedling substrate block. Discrete element analysis usually uses the EDEM software. Through EDEM simulation, the effect of grasping different viscous matrices

can be compared [14,57]. By utilizing the EDEM–RecurDyn coupling simulation [62], the force exerted on the gripper can be assessed while concurrently analyzing the effect of the clamping force on substrate particles.

3.1.3. Image Technology

Image technology can directly observe the internal structure of the substrate and study the mechanism of substrate damage at the microscopic level. The commonly used techniques include CT and electron microscopy.

The utilization of CT technology [38,64,65] to scan the root system of seedlings enables the generation of a three-dimensional reconstruction of the root system, allowing for observation of the growth and distribution of the root system. Additionally, the impact of diverse composite substrates on seedling growth quality and substrate strength can be investigated by using electron microscopy (SEM) images [66], providing a theoretical basis for exploring the mechanism of substrate damage.

3.2. End Effector

The end effector is the component, which directly contacts the substrate block in the tray. Its function is to grasp the seedlings from a tray and plant them in another tray or pot. Its structure directly affects the operation quality and efficiency of the transplanter. According to the movement form of the gripper, the end effector mainly includes three types: the plug-in and clamping type, oblique insertion type, and deformed sliding needle type. The gripper can be shaped like a shovel or needle; the number of gripper fingers is usually two to four shovels or needles; and the driving methods include cylinder, electric push rod, motor, and electromagnetic [49]. A summary comparison of the end effector structures is shown in Table 4.

The plug-in and clamping type end effector performs insertion and gripping actions during the operation and is widely used in transplanters. In order to achieve the action of inserting and grasping, the plug-in and clamping end effectors are usually controlled by two sets of drives: one for insertion and retraction of the needle, the other for opening and closing of the needle [68]. Some researchers have also designed a linkage mechanical structure with a return spring using a single cylinder to achieve the insertion and clamping actions of the end effector [69,70]. There are also some plug-in and clamping type end effectors, which are limited to gripping actions only and need external manipulation to complete the insertion process [38,71].

The oblique plug-in type end effector refers to a gripper, which is inserted into the substrate at an oblique angle along the hole. Due to the oblique arrangement of the needle, the drive mechanism needs to convert the vertical linear motion of the drive components into linear motion with a certain tilt angle. The oblique insertion mechanism can be a guiding groove mechanism [57,72] or a connecting rod mechanism [53]. The driving unit can be a cylinder [14,54,57] or an electric push rod [53].

Both of these methods will increase the size of the end effector and make it inconvenient to integrate multiple end effectors; however, the structure is reliable and operates stably.

The deformation sliding needle type end effector usually uses the flexible seedling needle and the arc-shaped guide tube to realize the inclined insertion of the seedling needle [73–76]. This structure is compact, but due to the friction between the flexible needle and the guide tube during the operation, the wear of the needle and the tube may occur during long-term use, and poor lubrication may cause the needles to jam, affecting the operating stability.

Table 4. Structure comparison of end effectors.

Structural Style	Drive	Gripper Shape	Finger/Shovel Quantity	Principle Characteristics	Reference
Plug-in clamping type	Single cylinder	Aciculiform	4	Insert and clamp the substrate continuously within one push stroke of the cylinder.	[69]
	Single cylinder	Aciculiform	4	The cylinder piston rod retracts during the process of gripper insertion into the substrate to achieve the grasping action.	[38]
	Single cylinder	Aciculiform	4	The crank slider mechanism is used to drive the gripper, and the end effector can realize two working modes of oblique insertion and clamping.	[70]
	Double cylinder and air sac	Aciculiform	4	Two cylinders with a certain angle are used to drive gripper insertion into the substrate obliquely. The airbag between the two cylinders and the tightening spring are used to realize the clamping and opening of the gripper.	[68]
Oblique plug-in type	Single cylinder	Spade shape	4	The finger shovel is driven by the air cylinder and inserted along the four walls of the hole to reduce the damage to the seedling mound. By adding blocks, the substrate on the shovel is removed during the shovel recovery.	[57]
	Single cylinder	Aciculiform	4	Using a single cylinder and the guide groove mechanism to realize oblique insertion action of the gripper.	[54]
	Linear effector	Aciculiform	4	Using electric push rod and the connecting rod mechanism to drive gripper insertion into the substrate.	[53]
	Single cylinder	Spade shape	4	Using one cylinder completes the extension and retraction action of four groups of seedling spades.	[72]
Deformation sliding needle type	Double cylinder	Aciculiform	4	The large cylinder drives the small cylinder and the needle fixing plate to move downward together. Then, the small cylinder pushes the needle fixing plate to move downward to realize the clamping action.	[73]
	Single cylinder	Aciculiform	4	The single cylinder pushes four flexible needles through four oblique guide tubes to grasp the seedling substrate.	[77]
	Servo motor	Aciculiform	4	When picking up seedlings, the four sliding needles extend out of the guide tube and are inserted obliquely into the hole to hold the seedling substrate.	[76]
	Single cylinder	Aciculiform	4	With the pushing force of the seedling claw control cylinder, the gripping action is achieved by the deformation of the seedling needles.	[75]

3.3. Integration for Multiple End Effectors

In order to improve the efficiency of transplanters, multiple end effectors are usually integrated to realize the grasping of multiple plants or entire rows of seedlings with just a single movement of the manipulator. According to different structures, the integration methods can be divided into the expanding and contracting board type, parallelogram mechanism, rope connection type, cam type, independent control type, etc.

The expanding and contracting board type is simple in structure and control; according to actual needs, the board can be arranged horizontally [78] or vertically [79], but this structure is heavy, which limits the action speed of the system [80]. The parallelogram mechanism type is a lightweight spacing adjustable mechanism for integrating multiple end effectors [81]. This mechanism is capable of achieving precise separation of all end effectors with minimal stroke [37].

In cases where the closing and opening distances between the multiple end effectors are fixed, a simpler method is to use a soft belt [77] or steel wire [82] to connect adjacent end effectors. As the soft belt or steel wire rope are flexible bodies with certain elasticity, they can play a buffering role in the process of distance change. Tong [77] used a cylinder to make the grippers expand and close. However, these methods are limited to a single specification of the tray; the length of the soft belt or steel wire must be readjusted if the tray specification changes. The cylindrical cam [83] mechanism can also be used for adjusting the distance for multiple end effectors, but this kind of structure requires high machining and installation accuracy.

The above mechanisms are simple in structure and easy to control, but they can only realize the equidistant expansion and closing of multiple end effectors, and most of these

can only be suitable for one kind of tray. If those kinds of mechanisms need to be applied to other hole spacings, the spacing between end effectors needs to be readjusted manually. In order to improve the versatility of the transplanter, many transplanter manufacturers have designed independent servo-driven end effectors, which can realize the adjustment of any distance between the end effectors.

The FlexPlanter and PackPlanter series transplanters produced by the TTA company employ independently controlled electric motors to drive the end effectors, which mesh with laterally arranged racks to achieve translational motion. The RW series transplanter produced by Urbinati also uses an independent motor to drive the end effector, only changing the transmission mechanism from gear rack to synchronous wheel and synchronous belt transmission. Feng et al. [40,84] have also developed a transplanter integrated with an end effector driven independently by synchronous wheel and synchronous belt.

3.4. Transplanting Manipulator

The function of a transplanting manipulator is to drive the end effector to move back and forth between the supply and target tray. There are four types of commonly used transplanting manipulators, including serial industrial robots, four-axis SCARA robots, parallel robots, and Cartesian coordinate manipulators. Table 5 summarizes the unique characteristics of transplanters with different structures.

Table 5. Comparison of transplanting manipulators.

Transplanting Manipulator	Structural Complexity	Bearable Load	DOF	Efficiency	Cost	Dimensions	Scalability
Serial Robot	Simple	High Load	5–6	High	Expensive	Middle	Best
Four-Axis SCARA Robot	Simple	Middle Load	4	Middle	Middle	Small	Good
Parallel Robot	Simple	Low Load	2–3	Low	Low	Small	Good
Cartesian Coordinate Manipulator	Complicated	High Load	2–3	High	High	Big	Bad

3.4.1. Series Industrial Manipulators

The earliest research on transplanters was generally based on industrial robotic arms [85], such as the Puma560 industrial robot [15] and the ABB robot [18]. These initial studies demonstrated the feasibility of using industrial robots for tray seedling transplanting operations, but they did not conduct practical application research. However, in recent years, there have been increased applications of industrial robot transplanters in vegetable production practice. For example, the Danish BEKIDAN company designed a flower grading transplanting system based on the DENSO six-axis industrial robot. Additionally, the Dutch horticultural production solution provider CODEMA company developed a transplant robot workstation using the KUKA robot, and the FANUC robot, equipped with a multiple end effectors' manipulator. Serial industrial robots provide a simple and space-efficient option for factory layout and installation, but they are more expensive.

3.4.2. SCARA Manipulator

The SCARA manipulator has become a popular choice for transplanting operations due to its compact structure, high speed, and high loads. Several researchers and companies have developed transplanting solutions using SCARA robots. Appropriate loads for SCARA robotic arms can be selected according to the requirements of transplanting operations, and they can be equipped with single [16,17], or multiple end effectors. However, SCARA robots typically have only four degrees of freedom, which limits their range of motion. To increase the motion range of the SCARA robot in transplantation operations, it is possible to install the robot on a linear module [86].

3.4.3. Parallel Manipulator

Parallel manipulators are known for their high rigidity and high speed. By using a two-degree-of-freedom parallel manipulator in combination with a conveyor, transplantation operations can be realized [87,88]. However, three-degree-of-freedom parallel manipulators are more commonly used [89]. The Italian Otechmek company, and Canadian CMP Automation Company [90], have both developed transplanting machines based on parallel manipulators. Due to the low load-bearing capacity of parallel manipulators, they are usually equipped with only a single end effector, which limits the efficiency of the entire transplantation operation.

3.4.4. Cartesian Coordinate Manipulator

The Cartesian coordinate manipulators have a simple and intuitive coordinate system, which makes them easy to program and operate. They are rigid and stable, which makes them suitable for heavy loads and high-speed operations. Therefore, this structure is often preferred in transplanters for vegetable seedlings. In Cartesian coordinate manipulators, there are two types of configurations available: two degrees of freedom and three degrees of freedom. In early research, the two-degree-of-freedom Cartesian coordinate manipulator was commonly used [19,20] due to its low cost and simple structure. However, the transplanters based on a three-degree-of-freedom Cartesian coordinate manipulator have a large workspace [40,91], as they can move in all three axes (x, y, z).

Due to the high load-bearing capacity of transplanters based on Cartesian coordinate manipulators, these machines are typically equipped with multiple end effectors to enhance operational efficiency. In this type of transplanter structure, there are various installation forms of end effectors according to the actual operational requirements, such as fixed installation and adjustable spacing installation. With fixed installation, multiple end effectors are closely adjacent to each other, and the spacing between them cannot be adjusted. The position of the end effectors must be adjusted by the translational axis of the Cartesian coordinate manipulator, which drives all the end effectors to move as a whole. Examples of such fixed installation end effectors include combifix II from the TTA company (TTA, Bleskensgraaf, The Netherlands) and Vision Planter from the ISO Group company (ISO-Horti Innovators, Gameren, The Netherlands). For the adjustable spacing installation, independent servo motors are typically used to drive the end effectors, such as the Pic-O-Mat Blueline, from the Viscon company (Viscon Group, Gravendeel, The Netherlands), and TEA 600N, from the Hamilton Design & TEA Project Company (Hamilton Design USA, Burton, OH, USA). When the end effectors can be driven independently, numerous cables or air pipes usually need to be arranged, such as the SPH-Transplanter, from the Flier Systems Company (Flier Systems, Barendrecht, The Netherlands), and the PlugPlanter S Model, from the Bouldin & Lawson company (Bouldin and Lawson, LLC, McMinnville, TN, USA). However, too many cables and organs can easily become entangled, leading to failures. In order to solve this problem, wireless end effectors have been developed, such as the PackPlanter Wireless transplanter, and the RW64 greenhouse transplanter, from the Italian Urbinati Company (Urbinati S.r.l., San Mauro Pascoli, Italy).

It can be observed that horticultural equipment manufacturing and solution providers around the world rely on their industrial base, and advanced technologies, such as machine vision, servo drive, and human machine interface (HMI), have developed various forms and functions of transplanters based on rectangular coordinate manipulators, which are widely used in facility horticulture. However, this large-scale transplanting equipment requires specialized hole trays, top seedling devices, etc., and its high cost hinders its promotion and application in small-scale facility agricultural environments. We summarize the key performance parameters of the commercial transplanters from various countries, as shown in Table 6.

Table 6. Commercial transplanters and technical parameters.

Country /Manufacturer	Model	Number of Grippers	Wired/Wireless	Independent Drive	Independent Clamp	Efficiency (Plants/Hour)	Applicable Plug Tray	Vision System	Air Consumption (L/min)	Power (kw)	Weight (kg)
The Netherlands/Viscon	Pic-O-Mat Blueline	4–8	Wired	Y	Y	10,000	Max size 600 × 400	N	--	--	--
	Pic-O-Mat Greenline	4–14	Wired	Y	Y	21,000	Max size 600 × 400	N	--	--	--
	Pic-O-Mat Redline	8–24	Wired	Y	Y	35,000	Max size 600 × 400	N	--	--	--
	Pic-O-Mat PC11	2–4	Wired	Y	Y	6000	4 plants per pot	N	--	--	--
	Pic-O-Mat PFS-8	8	Wired	Y	Y	10,000	4 plants per pot	N	--	--	--
	Pic-O-Mat VMP	6–12	Wired	Y	Y	12,000	4 plants per pot	N	--	--	--
	Fix-O-Mat TIFS-IV	12	Wired	Y	Y	12,000	--	Y	--	--	--
	Select-O-Mat Phoenix	12	Wired	Y	Y	8000	--	Y	--	--	--
The Netherlands/TTA	FlexPlanter	Multiple	Wired	Y	Y	3000–30,000	Plug size 9–30 mm	Y	60	5	2000
	FlexPlanter XF	Multiple	Wired	Y	Y	10,000–30,000	Plug size 9–30 mm	Y	15	--	4500
	PackPlanter wireless	Multiple	Wireless	Y	N	10,000–60,000	Plug size 9–30 mm	N	20	2.5	800
	PackPlanter	Multiple	Wireless	N	N	10,000–50,000	Plug size 9–30 mm	N	20	2	550
	PackPlanter S	Multiple	Wireless	Y	N	10,000–20,000	Plug size 9–30 mm	N	20	2	450
	MidiFlat	Multiple	Wireless	Y	N	4000–40,000	Plug size 9–60 mm	N	17	3	700
	MidiVision	Multiple	Wireless	Y	Y	5000–40,000	Plug size 9–60 mm	Y	17	3.5	800
	FlexSorter	Multiple	Wired	Y	Y	3000–12,000	Plug size 9–60 mm	Y	60	5	2000
	FlexSorter XF	Multiple	Wired	Y	Y	10,000–30,000	--	Y	16	--	4500
	MaxSorter	Multiple	Wired	Y	Y	6000–12,000	--	Y	35	--	3800
	Combifix II	Multiple	Wired	N	Y	12,000–20,000	Plug size 9–30 mm	Y	535	5	2250
The Netherlands/Flier Systems	SPH-Transplanter	Multiple	Wired	Y	Y	--	--	Y	--	--	--
	Young Plant Sorter	Multiple	Wired	N	Y	8000	--	Y	--	--	--
	Plug Fixer	5	Wired	N	Y	11,000	--	Y	--	--	--
The Netherlands/ISO Group	Vision Planter	Multiple	Wired	N	Y	--	--	Y	--	--	--
	Plug Planting Machine	Multiple	Wired	N	Y	--	--	N	--	--	--
Great Britain/TEA	TEA 600N	12	Wired	Y	Y	18,000	Plug size 30–50 mm	N	70	1	600
	TEA 1500N	12	Wired	Y	Y	12,000–14,400	Max size 600 × 400	N	80	1	700
	TEA 2000N	16	Wired	Y	Y	16,000	Max size 600 × 400	N	80	1	750
	TEA 1500J	2–12	Wired	Y	Y	12,000–14,400	Max size 600 × 400	N	80	1	650
	TEA 2000J	16	Wired	Y	Y	16,000	Max size 600 × 400	N	80	1	700
Italy/Urbinati	RW32	40	Wireless	Y	N	40,000	--	--	80	4	960
	RW64	80	Wireless	Y	N	56,000	--	--	80	5	1300
USA/Bouldin & Lawson	PlugPlanter S Model	16–32	Wired	Y	N	16,400–32,500	--	--	--	--	--

3.5. Growth Status Identification for Seedlings

Among the four types of transplanting operations, the widening-spacing, as shown in Figure 1b, and grading transplanting, as shown in Figure 1c, both require identification of the growth status of seedlings. The system for recognizing healthy seedlings is summarized in Table 7.

Leaf area is one of the most suitable characteristics of seedlings used in evaluating quality. Researchers have studied the use of machine vision technology in the measurement

of the leaf area [92]. Generally, the leaf images are obtained by a monocular CCD camera, and the leaf area is obtained using a threshold segmentation method [92–94]. Leaf color, leaf number, and plant height [95] can also serve as parameters for evaluating the growth of plug tray seedlings. The plant height data can be obtained using a line laser [94,95] or depth camera [96,97].

The feature extraction of seedling characteristics in the above methods is based on traditional machine vision morphological methods. Such methods are greatly affected by image background, lighting, and shadow of occlusion, and the stability of recognition performance is not ideal [98,99]. With the development of computer power and deep neural networks, the image recognition technology based on convolutional neural networks [100,101] has been increasingly applied in the identification of healthy seedlings for transplanting machines.

By using binocular vision [102] or depth cameras [97,103,104] to obtain point clouds of seedlings, three-dimensional reconstruction techniques can be employed to obtain the three-dimensional model of the seedlings. This method can provide more growth information, thereby improving the accuracy of identification and grading for plug tray seedlings.

Table 7. Vision identification systems.

Reference	Method	Camera Model	Resolution	Light Source	Collected Information	Algorithm	Performance
[94]	Monocular CCD camera, line laser	--	--	650 nm wavelength red light source line laser	Leaf area and plant height	Thresholding and linear structured light 3D location algorithm	Plant height accuracy is 5 mm.
[92]	Monocular CCD camera	PU LNIX TMC-7DSP	640 × 480	6 fluorescent lamps (380–780 nm), 40 W	Leaf area	Improved watershed algorithm and threshold segmentation	The recognition accuracy is 98%, and the average recognition time of a disc of seedlings is 4.396 s.
[95]	Monocular CCD camera, line laser	PULNiX TM-7CN	768 × 484	770–790 nm laser, 2 quartz halogen light sources, 300 W	Leaf area and plant height	Template matching and line structured light 3D positioning algorithm	The accuracy of hole identification is 95%.
[93]	Monocular camera	Pulnix TM-745	--	2 tungsten filament lamps, 120 W	Leaf area	Otus adaptive threshold segmentation	--
[19]	Monocular CCD camera	CCD	--	--	Leaf area	Threshold segmentation	--
[105,106]	Monocular camera	JoinHope Image OK-AC1300	--	Four F40BX/480 fluorescent lamps, 36 W	Side view of seedlings at 0 and 90 positions (upright degree and plant height)	Threshold segmentation	Each image processing algorithm takes 0.35 s on average.
[102]	3D stereo positioning with binocular vision	DaHeng Image DH-GV400UC	752 × 480	--	Top view of hole seedling	SIFT feature matching algorithm	The three-dimensional reconstruction image of the original acupoint plate is obtained.
[107]	Monocular camera	Epson Inc. GT800	512 × 512	--	Seedling hole area and total area	Three-layer neural network algorithm	--
[108]	Monocular camera	JHSM300E	3 megapixel	--	Top view of hole seedling	Threshold segmentation	The accuracy of hole identification is 100%.
[97]	Depth camera	Intel RealSense SR300 depth sensor	640 × 480	--	Seedling depth image	Point cloud clustering algorithm	The identification accuracy is 96.59% for 105-hole disc of seedlings with 10 days growth time.
[109]	Monocular CCD camera	PU LNIX TMC-7DSP	640 × 480	Six F40BX/480 fluorescent lamps, 36 W	Leaf area and blade circumference	Improved watershed algorithm and threshold segmentation	The identification accuracy of inferior seedlings is above 98%.
[96]	Depth camera	Realsense D415	1280 × 720	30 mm × 90 mm kM-BRD7530 LED	Seedlings' height and edge points	ExG algorithm	Image processing average time is 0.753 s.
[110]	Monocular CCD camera	DaHeng MER-131-75GC	1280 × 1024	Lemons OPT-L133037-W	Plug seedlings, plug bodies, and stem leaves	Threshold segmentation, morphological processing	--

3.6. Path Planning Methods

The transplanting operation of plug tray seedlings requires the manipulator to move back and forth between the supply tray and the target tray. The length of the manipulator's moving path is different with different moving strategies, which affects the operation efficiency of the transplanter.

In the path planning methods of the transplanter, the simplest kind is the fixed order method, such as the far-to-near, near-to-far, or snake-like trajectory [111], where the sequence of seedling selection is fixed, and real-time calculation of the path is not required, but the path distance is generally longer, and the efficiency is lower. Other methods are based on common optimization algorithms, such as the ant colony algorithm [112], genetic algorithm [113], greedy algorithm, etc. Some scholars have also combined and improved several algorithms [114] for trajectory planning. The trajectory planning methods based on optimization algorithms can optimize the path distances, but they take a longer time. Therefore, it is recommended to use optimization algorithms for transplanter trajectory planning after comparison with the fixed order methods.

Overall, the path planning strategy for vegetable seedling transplanting should consider the tray layout, space availability, and operation efficiency. By carefully planning the manipulator's track, the transplanter can improve the efficiency of the transplanting operation and reduce energy costs.

4. Discussion

4.1. Summary of Current Problems

By comparing the research status of plug tray seedling transplanters in different countries and regions, significant disparities in the technological level and application of transplanting machines can be observed. The significant gaps between developed and developing countries can be attributed to several factors, such as economic resources, infrastructure and agricultural development, technological knowledge level, policy and government support. Based on the above reasons, there are differences in the production scale, planting mode, planting density, hardware configuration, and other aspects of protected agriculture in different countries and regions. Transplanters have limitations in intelligence, applicability, and universality. In order to promote better application of transplanters in various protected agricultural conditions, the following problems need to be resolved in future research:

- Limited standardization and compatibility. In protected agriculture, variations exist in the dimensions of seedbeds, planting densities, and tray specifications utilized by different regions and growers. Furthermore, there is a wide range of seedling substrate compositions and proportions employed. Manual seeding predominates, leading to inconsistencies in seedling position. Additionally, the timing of transplanting operations is not standardized. All these differences in the production modes of vegetables have led to the poor practicability of transplanters. Therefore, it is necessary to standardize the production mode by unifying the supporting planting equipment and planting agronomy, combining agronomy with equipment, so as to reduce the complexity of transplanting equipment development.
- Insufficient intelligence and automation. The existing transplanting machines have made some progress in terms of intelligence, but there is still room for improvement. Some aspects include sensing and perception, decision making and control, adaptability, learning, and human–machine interaction. With the development of artificial intelligence technology, we can develop a plant growth status recognition system based on deep learning and three-dimensional reconstruction technology, conduct research on motion control algorithms of multiple grippers to improve the transplanting efficiency, and use the internet of things (IoT) technology to develop remote monitoring functions for transplanters.
- Limited adaptability to diverse agricultural environments. Transplanting machines often face challenges in adapting to the wide range of agricultural environments

found in different regions and countries. Variations in climate, soil conditions, crop varieties, and planting practices necessitate flexible and adaptable machines. Currently, there is a lack of transplanting machines, which can easily accommodate these variations, resulting in suboptimal performance and reduced efficiency. Future research should focus on developing transplanting machines, which can be easily customized and configured to suit different agricultural environments, ensuring optimal transplanting outcomes.

- Cost effectiveness and affordability. The high cost of transplanting machines is a significant barrier to their widespread adoption, particularly in developing countries and small-scale farming operations. The extensive equipment sets and specialized components required for transplanting machines contribute to their high procurement and maintenance costs. To encourage broader adoption, research efforts should concentrate on developing cost-effective solutions, including the use of affordable materials, streamlined designs, and modular components. By reducing the overall cost of transplanting machines, their accessibility and affordability can be enhanced, facilitating their integration into diverse agricultural systems.

4.2. Research Focus and Development Trend

Considering the aforementioned issues associated with transplanters in protected agriculture, the following research directions are recommended to be prioritized for further investigation:

- Integration of advanced sensing technologies. One of the prominent trends in transplanting machine research is the integration of advanced sensing technologies. This includes the incorporation of diverse sensors, such as cameras, 3D scanners, lidar, thermal imagers, and hyperspectral sensors. These sensors enable precise and comprehensive data acquisition, facilitating a deeper understanding of both the crop and the surrounding environment. By integrating data from multiple sensors, transplanting machines can enhance their perception capabilities, enabling real-time monitoring and analysis of crucial parameters, such as soil moisture, plant health, and environmental conditions.
- Development of intelligent decision-making algorithms. Another key focus in transplanting machine research is the development of intelligent decision-making algorithms. Machine-learning techniques, such as deep learning, can be applied to analyze the collected data and extract meaningful insights. These algorithms enable transplanting machines to make informed decisions based on real-time information, optimizing transplanting strategies and enhancing overall efficiency. Additionally, the integration of artificial intelligence and machine vision technologies allows for automated detection and classification of seedlings, improving the accuracy and precision of transplanting operations.
- Advancements in robotics and automation. The advancement of robotics and automation technologies plays a significant role in the evolution of transplanting machines. Researchers are exploring the development of robotic systems with improved dexterity, allowing for more precise and efficient handling of seedlings. Automation features, such as autonomous navigation, adaptive grasping, and coordinated multi-robot systems, are being investigated to enhance the performance and productivity of transplanting machines. These advancements aim to reduce the reliance on human labor, increase operational efficiency, and minimize human errors.
- Compact and lightweight technology for transplanting machines. The development of compact and lightweight technology for transplanting machines offers a promising solution to address the challenges related to the adaptability and cost effectiveness of large-scale equipment. By adopting smaller and more flexible transplanter systems, greater versatility can be achieved, allowing for tailored configurations and programming specific to different crop types and environmental conditions. The utilization of finite element analysis techniques enables the optimization of structural design for

transplanting machine components, while incorporating lightweight materials in the manufacturing process enhances overall performance. Furthermore, the downsizing of equipment also contributes to cost reduction. Integrating cost reduction into this discussion, the development of compact and lightweight transplanting machines not only enhances adaptability but also improves cost effectiveness, making them more accessible to a wider range of users.

5. Conclusions

This review highlights the existing gaps and challenges in the applicability of current transplanters to diverse protected agricultural environments. To address these issues, the development of compact general-purpose transplanters is recommended to promote their widespread adoption in various protected agricultural settings. However, the variations in planting modes and hardware configurations of plug seedlings present a significant challenge in achieving compact universal transplanters.

As agricultural labor declines and protected agriculture continues to advance, the mechanization and automation of protected agriculture production become inevitable. To overcome the associated difficulties and challenges, it is crucial to closely monitor technological advancements in the industry and develop unified standards for protected agriculture planting. A collaborative approach involving experts from diverse disciplines, particularly engineers and researchers specializing in protected agriculture planting, is essential. By fostering the synergistic integration of agricultural mechanization and agronomic principles, alongside the standardization of protected vegetable cultivation, the widespread adoption and utilization of automated plug tray seedling transplanting equipment can be effectively facilitated. This transformation holds the potential to enhance the applicability and efficacy of transplanting machines in protected agriculture, resulting in improved efficiency and cost reduction. It requires continuous research and development efforts to address the specific needs and challenges of different agricultural environments. Furthermore, interdisciplinary collaborations and knowledge exchange are vital for driving innovation and advancing the field of transplanting machine technology. By leveraging these strategies, the vision of efficient and sustainable protected agriculture can be realized, contributing to the growth and development of modern agricultural practices.

Author Contributions: W.L., Conceptualization, Investigation, Writing—original draft, Visualization; S.T., Writing—review and editing; Q.W., Writing—review and editing; H.J., Conceptualization, Supervision, Project administration, Writing—review and editing. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by the Science and Technology Program of Zhejiang Province, grant No. 2022C02026 and No. 2022C02039, China.

Institutional Review Board Statement: Not applicable.

Data Availability Statement: No new data were created or analyzed in this study. Data sharing is not applicable to this article.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Buturi, C.V.; Mauro, R.P.; Fogliano, V.; Leonardi, C.; Giuffrida, F. Mineral Biofortification of Vegetables as a Tool to Improve Human Diet. *Foods* **2021**, *10*, 223. [CrossRef] [PubMed]
2. Ali, A.; Ghani, M.I.; Haiyan, D.; Iqbal, M.; Cheng, Z.; Cai, Z. Garlic Substrate Induces Cucumber Growth Development and Decreases Fusarium Wilt through Regulation of Soil Microbial Community Structure and Diversity in Replanted Disturbed Soil. *Int. J. Mol. Sci.* **2020**, *21*, 6008. [CrossRef] [PubMed]
3. Guo, D.; Juan, J.; Chang, L.; Zhang, J.; Huang, D. Discrimination of Plant Root Zone Water Status in Greenhouse Production Based on Phenotyping and Machine Learning Techniques. *Sci. Rep.* **2017**, *7*, 8303. [CrossRef] [PubMed]
4. Kang, M.H.; Jeon, S.S.; Shin, S.M.; Veerana, M.; Ji, S.-H.; Uhm, H.-S.; Choi, E.-H.; Shin, J.H.; Park, G. Dynamics of Nitric Oxide Level in Liquids Treated with Microwave Plasma-Generated Gas and Their Effects on Spinach Development. *Sci. Rep.* **2019**, *9*, 1011. [CrossRef] [PubMed]

5. Maher, A.; Kamel, E.; Enrico, F.; Atif, I.; Abdelkader, M. An Intelligent System for the Climate Control and Energy Savings in Agricultural Greenhouses. *Energy Effic.* **2016**, *9*, 1241–1255. [CrossRef]
6. Shi, X.; An, X.; Zhao, Q.; Liu, H.; Xia, L.; Sun, X.; Guo, Y. State-of-the-Art Internet of Things in Protected Agriculture. *Sensors* **2019**, *19*, 1833. [CrossRef]
7. Hiwasa-Tanase, K.; Ezura, H. Molecular Breeding to Create Optimized Crops: From Genetic Manipulation to Potential Applications in Plant Factories. *Front. Plant Sci.* **2016**, *7*, 539. [CrossRef]
8. Li, J.; Wu, T.; Huang, K.; Liu, Y.; Liu, M.; Wang, J. Effect of LED Spectrum on the Quality and Nitrogen Metabolism of Lettuce Under Recycled Hydroponics. *Front. Plant Sci.* **2021**, *12*, 678197. [CrossRef]
9. Patil, J.A.; Yadav, S.; Kumar, A. Management of Root-Knot Nematode, Meloidogyne Incognita and Soil Borne Fungus, Fusarium Oxysporum in Cucumber Using Three Bioagents under Polyhouse Conditions. *Saudi J. Biol. Sci.* **2021**, *28*, 7006–7011. [CrossRef]
10. Viršilė, A.; Brazaitytė, A.; Vaštakaitė-Kairienė, V.; Miliauskienė, J.; Jankauskienė, J.; Novičkovas, A.; Laužikė, K.; Samuolienė, G. The Distinct Impact of Multi-Color LED Light on Nitrate, Amino Acid, Soluble Sugar and Organic Acid Contents in Red and Green Leaf Lettuce Cultivated in Controlled Environment. *Food Chem.* **2020**, *310*, 125799. [CrossRef]
11. Zou, T.; Huang, C.; Wu, P.; Ge, L.; Xu, Y. Optimization of Artificial Light for Spinach Growth in Plant Factory Based on Orthogonal Test. *Plants* **2020**, *9*, 490. [CrossRef] [PubMed]
12. Hanafi, A.; Papasolomontos, A. Integrated Production and Protection under Protected Cultivation in the Mediterranean Region. *Biotechnol. Adv.* **1999**, *17*, 183–203. [CrossRef] [PubMed]
13. Tong, J.; Qiu, Z.; Zhou, H.; Bashir, M.K.; Yu, G.; Wu, C.; Du, X. Optimizing the Path of Seedling Transplanting with Multi-End Effectors by Using an Improved Greedy Annealing Algorithm. *Comput. Electron. Agric.* **2022**, *201*, 107276. [CrossRef]
14. Tian, Z.; Ma, W.; Yang, Q.; Yao, S.; Guo, X.; Duan, F. Design and Experiment of Gripper for Greenhouse Plug Seedling Transplanting Based on EDM. *Agronomy* **2022**, *12*, 1487. [CrossRef]
15. Kutz, L.J.; Miles, G.E.; Hammer, P.A.; Krutz, G.W. Robotic Transplanting of Bedding Plants. *Trans. Asae* **1987**, *30*, 586–590. [CrossRef]
16. Tmg, K.C.; Giacomelli, G.A.; Shen, S.J. Robot Workcell for Transplanting of Seedlings Part I—Layout and Materials Flow. *Trans. Asae* **1990**, *33*, 1005–1010. [CrossRef]
17. Ting, K.C.; Giacomelli, G.A.; Shen, S.J.; Kabala, W.P. Robot Workcell for Transplanting of Seedlings. Part II. End-Effector Development. *Trans. Asae* **1990**, *33*, 1013–1017. [CrossRef]
18. Hůla, P.; Šindelář, R.; Trinkl, A. Verification of Applicability of ABB Robots for Trans-Planting Seedlings in Greenhouses. *Res. Agric. Eng.* **2008**, *54*, 155–162. [CrossRef]
19. Ryu, K.H.; Kim, G.; Han, J.S. AE—Automation and Emerging Technologies. *J. Agric. Eng. Res.* **2001**, *78*, 141–146. [CrossRef]
20. Kang, D.H.; Kim, D.E.; Lee, G.I.; Kim, Y.H.; Min, Y.B. Development of a Vegetable Transplanting Robot. *J. Biosyst. Eng.* **2012**, *37*, 201–208. [CrossRef]
21. Onosaka, T.; Okuno, K.; Uchida, K.; Buno, S.; Yamada, H. Transplanter. U.S. Patent 5842306A, 22 May 1996.
22. Kubota IKP-4 Vegetable Tranplanter. Available online: https://www.kubota.com.cn/kams/productlist.do?method=list&&modelCodeNow=104 (accessed on 15 July 2023).
23. Yanmar PF2R Vegetable Tranplanter. Available online: https://www.yanmar-china.com/cn/agri/products/vegetable_replant/pf2r/ (accessed on 15 July 2023).
24. GFM Planting. Available online: https://www.gfm.co.jp/product/detail/gfm_planting.html (accessed on 15 July 2023).
25. GFM Transplanter. Available online: https://www.gfm.co.jp/product/detail/gfm_transplant.html (accessed on 15 July 2023).
26. Transplanting Machines for Young Plants. Available online: https://viscongroup.eu/machines/transplanting-machines/ (accessed on 11 February 2023).
27. TTA Offers a Wide Range of Transplanting Equipment. Available online: https://www.tta.eu/equipment/transplanting (accessed on 12 February 2023).
28. Flier Systems Helps Ambitious Growers Flourish—Flier Systems. Available online: https://fliersystems.com/en/ (accessed on 12 February 2023).
29. Machines—ISO Group Agri Systems BV. Available online: https://www.iso-group.nl/en/machines (accessed on 12 February 2023).
30. Codema Systems Group. Available online: https://vimeo.com/codemasystems (accessed on 12 February 2023).
31. Transplanting—Nurseries Transplanting Machines. Available online: https://www.urbinati.com/en/applications/transplanting-machines-plant-nursery/ (accessed on 12 February 2023).
32. Teaproject. Available online: http://www.teaproject.it/transplanters.htm (accessed on 12 February 2023).
33. Transplanting Systems—Agrinomix. Available online: https://agrinomix.com/transplanting-systems/ (accessed on 12 February 2023).
34. PlugPlanter S Model—Bouldin & Lawson. Available online: https://bouldinlawson.com/plugplantersmodel/ (accessed on 12 February 2023).
35. Demtec Dempic III. Available online: https://www.transplantsystems.co.nz/products/dempic/ (accessed on 12 February 2023).
36. Tian, Z.; Ma, W.; Yang, Q.; Yao, S.; Zhang, M.; Duan, F.; Xu, H. Research Status and Problem Analysis of Plug Seedling Transplanter in Greenhouse. *J. China Agric. Univ.* **2022**, *27*, 22–38. [CrossRef]

37. Yang, Q.; Xu, L.; Shi, X.; Ibrar, A.; Mao, H.; Hu, J.; Han, L. Design of Seedlings Separation Device with Reciprocating Movement Seedling Cups and Its Controlling System of the Full-Automatic Plug Seedling Transplanter. *Comput. Electron. Agric.* **2018**, *147*, 131–145. [CrossRef]
38. Mao, H.; Liu, Y.; Han, L.; Sheng, B.; Ma, G.; Li, Y. X-ray Computerized Tomography for Characterization of Pick-up Destruction and Pick-up Parameter Optimization of Tomato Root Lumps. *Span. J. Agric. Res.* **2019**, *17*, e0202. [CrossRef]
39. Han, L.; Kumi, F.; Mao, H.; Hu, J. Design and Tests of a Multi-Pin Flexible Seedling Pick-up Gripper for Automatic Transplanting. *Appl. Eng. Agric.* **2019**, *35*, 949–957. [CrossRef]
40. Feng, Q.; Xiu, W.; Zhao, C.; Kai, J.; Fan, P. Design and Test of Tray-Seedling Sorting Transplanter. *Biol. Eng.* **2015**, *8*, 7. [CrossRef]
41. Zhao, S.; Lei, X.; Liu, J.; Jin, Y.; Bai, Z.; Yi, Z.; Liu, J. Transient Multi-Indicator Detection for Seedling Sorting in High-Speed Transplanting Based on a Lightweight Model. *Comput. Electron. Agric.* **2023**, *211*, 107996. [CrossRef]
42. Khadatkar, A.; Pandirwar, A.P.; Paradkar, V. Design, Development and Application of a Compact Robotic Transplanter with Automatic Seedling Picking Mechanism for Plug-Type Seedlings. *Sci. Rep.* **2023**, *13*, 1883. [CrossRef]
43. Khadatkar, A.; Mathur, S.M. Design and Development of an Automatic Vegetable Transplanter Using a Novel Rotating Finger Device with Push-Type Mechanism for Plug Seedlings. *Int. J. Veg. Sci.* **2022**, *28*, 121–131. [CrossRef]
44. Dihingia, P.C.; Prasanna Kumar, G.V.; Sarma, P.K. Development of a Hopper-Type Planting Device for a Walk-Behind Hand-Tractor-Powered Vegetable Transplanter. *J. Biosyst. Eng.* **2016**, *41*, 21–33. [CrossRef]
45. Miah, M.S.; Rahman, M.M.; Hoque, M.A.; Ibrahim, S.M.; Sultan, M.; Shamshiri, R.R.; Ucgul, M.; Hasan, M.; Barna, T.N. Design and Evaluation of a Power Tiller Vegetable Seedling Transplanter with Dibbler and Furrow Type. *Heliyon* **2023**, *9*, e17827. [CrossRef]
46. Jorg, O.J.; Sportelli, M.; Fontanelli, M.; Frasconi, C.; Raffaelli, M.; Fantoni, G. Design, Development and Testing of Feeding Grippers for Vegetable Plug Transplanters. *AgriEngineering* **2021**, *3*, 669–680. [CrossRef]
47. Pei, D.; Meng, F.; Wang, H. Research Progress of Visual Inspection of Tray Seedling and the System of Automatic Transplanting. *Int. J. Multimed. Ubiquitous Eng.* **2016**, *11*, 57–68. [CrossRef]
48. Wen, Y.; Zhang, J.; Yuan, T.; Tan, Y. Current Situation and Analysis of Automatic Pick-up Technology for Vegetable Plug Seedlings. *J. China Agric. Univ.* **2021**, *26*, 128–142. [CrossRef]
49. Liu, W.; Liu, J. Review of End-Effectors in Tray Seedlings Transplanting Robot. *J. Agric. Mech. Res.* **2013**, *35*, 6–10. [CrossRef]
50. Wang, N.; Ren, L.; Li, J.; Hu, H.; Wang, B.; Shang, J. Research Status and Prospect of Automatic Seedling Picking Technology of Transplanter. *J. Chin. Agric. Mech.* **2021**, *42*, 59–66. [CrossRef]
51. Chen, D. Summary of Foreign and Domestic Vegetable Corps Plug Transplants Production. *China Veg.* **2000**, *1*, 9–13. [CrossRef]
52. Mohamed, S.; Liu, J. Effect of Soil Moisture Content and End-Effector Speed on Pick-up Force and Lump Damage for Seedling Transplanting. *AgriEngineering* **2019**, *1*, 343–356. [CrossRef]
53. Liu, J.; Li, M.; Li, N.; Li, P.; Zhao, M.; Yue, W. Design and Test of End-Effector for Automatic Transplanting of Strawberry Plug Seedlings. *Trans. Chin. Soc. Agric. Mach.* **2016**, *47*, 49–58. [CrossRef]
54. Gao, G.H.; Feng, T.; Li, F. Working Parameters Optimization and Experimental Verification of Inclined-Inserting Transplanting Manipulator for Plug Seedling. *Trans. Chin. Soc. Agric. Eng.* **2015**, *31*, 16–22. [CrossRef]
55. Tong, J.; Jiang, H.; Jiang, Z.; Cui, D. Experiment on Parameter Optimization of Gripper Needles Clamping Seedling Plug for Automatic Transplanter. *Trans. Chin. Soc. Agric. Eng.* **2014**, *30*, 8–16. [CrossRef]
56. Wang, M.; Song, J.; Liu, C.; Wang, Y.; Sun, Y. Design and Experiment of Crank Rocker Type Clamp Seedlings Mechanism of Vegetable Transplanter. *Trans. Chin. Soc. Agric. Eng.* **2015**, *31*, 49–57. [CrossRef]
57. Tong, J.; Shi, H.; Wu, C.; Ding, Y.; Zhao, X.; Wang, R. Simulation and Test of Seedling Pot Grabbing by Spade End-Effector. *Trans. Chin. Soc. Agric. Mach.* **2019**, *50*, 107–116. [CrossRef]
58. Miao, X.; Mao, H.; Han, L.; Sun, H.; Yang, X.; Huang, H. Analysis of Influencing Factors on Force of Picking Plug Seedlings and Pressure Resistance of Plug Seedlings. *Trans. Chin. Soc. Agric. Mach.* **2013**, *44*, 27–32. [CrossRef]
59. Han, L.; Mao, H.; Hu, J.; Miao, X.; Tian, K.; Yang, X. Experiment on Mechanical Property of Seedling Pot for Automatic Transplanter. *Trans. Chin. Soc. Agric. Eng.* **2013**, *29*, 24–29.
60. Han, L.; Zhao, Z.; Ma, G.; Ye, M.; Mao, H.; Hu, J. Compression-Force Relaxation Characteristics of Vegetable Pot. *Jiangsu Agric. Sci.* **2018**, *46*, 271–274. [CrossRef]
61. Wang, C.; Liu, C.; Li, Y.; Song, J.; Wang, J.; Dong, X. Design and Experiment of Pneumatic Punching High-Speed Seedling Picking Device for Vegetable Transplanter. *Trans. Chin. Soc. Agric. Mach.* **2021**, *52*, 35–43+51. [CrossRef]
62. Hu, J.; Pan, J.; Chen, F.; Yue, R.; Yao, M.; Li, J. Simulation Optimization and Experiment of Finger-Clamping Seedling Picking Claw Based on EDEM-RecurDyn. *Trans. Chin. Soc. Agric. Mach.* **2022**, *53*, 75–85+301. [CrossRef]
63. Sun, G.; Wang, X.; He, G.; Zhou, T.; Wang, C.; Qiao, X. Design of the End-Effector for Plug Seedlings Transplanter and Analysis on Virtual Prototype. *Trans. Chin. Soc. Agric. Mach.* **2010**, *41*, 48–53+47.
64. Tracy, S.R.; Black, C.R.; Roberts, J.A.; Sturrock, C.; Mairhofer, S.; Craigon, J.; Mooney, S.J. Quantifying the Impact of Soil Compaction on Root System Architecture in Tomato (Solanum Lycopersicum) by X-Ray Micro-Computed Tomography. *Ann. Bot.* **2012**, *110*, 511–519. [CrossRef]
65. Liu, Y.; Mao, H.; Han, L.; Xu, J.; Ma, G.; Li, Y. Plug Damage Detection and Parameter Optimization of Picking up Cucumber Seedlings from Tray Cells Based on Micro-CT. *Trans. Chin. Soc. Agric. Eng.* **2018**, *34*, 27–34. [CrossRef]

66. Han, L.; Mo, M.; Gao, Y.; Ma, H.; Xiang, D.; Ma, G.; Mao, H. Effects of New Compounds into Substrates on Seedling Qualities for Efficient Transplanting. *Agronomy* **2022**, *12*, 983. [CrossRef]
67. Gao, G.H.; Yang, H.; Wei, K.C. Simulation and Optimization Design of the Transplanting Manipulator. *Adv. Mater. Res.* **2013**, *712*, 2308–2311. [CrossRef]
68. Han, L.; Mao, H.; Yan, L.; Hu, J.; Huang, W.; Huang, L. Pincette-Type End-Effector Using Two Fingers and Four Pins for Picking Up Seedlings. *Trans. Chin. Soc. Agric. Mach.* **2015**, *46*, 23–30. [CrossRef]
69. Jiang, Z.; Jiang, H.; Tong, J. Optimal Design of End-Effector on Automatic Plug Seedling Transplanter. *J. Zhejiang Univ. Sci.* **2017**, *51*, 1119–1125. [CrossRef]
70. Vera, S.; Prado, S. Synthesis and Optimization of a Needles Robotic Gripper Mechanism for Transplanting Seedlings. In Proceedings of the 2020 IEEE Andescon, Quito, Ecuador, 13–16 October 2020; IEEE: Quito, Ecuador, 2020; pp. 1–6.
71. Yang, Q.; Huang, G.; Shi, X.; He, M.; Ahmad, I.; Zhao, X.; Addy, M. Design of a Control System for a Mini-Automatic Transplanting Machine of Plug Seedling. *Comput. Electron. Agric.* **2020**, *169*, 105226. [CrossRef]
72. Xu, G.; Xu, Y.; Zeng, J.; Cui, Y.; Zhao, Y.; Xiao, J.; Ningbo Institute of Materials Technology and Engineering Chinese Academy of Sciences. End Actuator and Method for Removing and Repairing Seedlings in Plug Tray. CN Patent 202210033570.5, 2022.
73. Gao, G.H.; Wei, K.C.; Yang, H.; Liu, Y. Optimization Design of Automatic Transplanting Machine Based on TRIZ. *Adv. Mater. Res.* **2013**, *712*, 2966–2969. [CrossRef]
74. Arirang Culture Arirang Special—M60Ep239C02 Agricultural Smart Technology System in Korea. Available online: https://www.youtube.com/watch?v=GP3nJY3nDVI (accessed on 22 February 2023).
75. Zhang, K.; Wu, H.; Wang, W.; Song, C.; Liu, X. Design and Experiment of Clamping Picking Seedling Mechanism for Tomato Transplanter. *J. Agric. Mech. Res.* **2020**, *42*, 64–68. [CrossRef]
76. Tian, S.; Qiu, L.; Kondo, N.; Yuan, T. Development of Automatic Transplanter for Plug Seedling. *IFAC Proc. Vol.* **2010**, *43*, 79–82. [CrossRef]
77. Tong, J.; Meng, Q.; Gu, S.; Wu, C.; Ma, K. Design and Experiment of High-Speed Sparse Transplanting Mechanism for Hydroponics Pot Seedlings in Greenhouses. *Trans. Chin. Soc. Agric. Eng.* **2021**, *37*, 1–9. [CrossRef]
78. Yan, L.; Mao, H.; Han, L.; Hu, J.; Wang, L. Design Ang Test of Motion Control Device for Tray Seedings Multi-Endingtransplanter in Greenhouse. *J. Agric. Mech. Res.* **2017**, *39*, 83–87. [CrossRef]
79. Zhu, C.; Li, W.; Gu, X.; Xie, D.; Li, C.; Zhang, Y. Analysis and Experiment of the Design and Test System of the Automatic Transplanting Equipment for Hole Plate Seedlings. *Agric. Eng. Technol.* **2021**, *41*, 23–29. [CrossRef]
80. Zhao, S.; Liu, J.; Zhou, X. Design and Experiment of Portable Greenhouse Transplanting Machine. *J. Agric. Mech. Res.* **2020**, *42*, 51–57. [CrossRef]
81. Assal, S.; Ndawula, I. *Optimum Design and FEA of a Hybrid Parallel-Deployable Structure-Based 3-DOF Multi-Gripper Translational Robot for Field Pot Seedlings Transplanting*; SCITEPRESS—Science and Technology Publications: Prague, Czech Republic, 2019; pp. 68–77.
82. Zhao, X.; Yang, Q.; Huang, G.; Hao, M.; Mao, H. Design and Test of Picking Seedling Mechanism of Small Full-Automatic Transplanter for Plug Seedlings. *J. Jiangsu Univ. Sci. Ed.* **2022**, *43*, 54–61. [CrossRef]
83. Cui, Y.; Wei, Y.; Ding, X.; Cui, G.; He, Z.; Wang, M. Design and Experiment of Adjustable Spacing End-Effector Based on Cylindrical Cam. *Trans. Chin. Soc. Agric. Mach.* **2022**, *53*, 104–114+122.
84. Feng, Q.; Wang, X.; Jiang, K.; Zhou, J.; Zhang, R.; Ma, W. Design and Test of Key Parts on Automatic Transplanter for Flower Seedling. *Trans. Chin. Soc. Agric. Eng.* **2013**, *29*, 21–27.
85. Hwang, H.; Sistler, F.E. A Robotic Pepper Transplanter. *Appl. Eng. Agric.* **1986**, *2*, 2–5. [CrossRef]
86. Liu, J.; Zhou, X. Full-Automatic System of Integrated Sorting-Transplanting-Replanting and Operation Method for Plug Seedlings. U.S. Patent 334,453, 27 August 2020.
87. Ma, J.; Hu, J.P.; Yan, X.Y.; Qi, C.H.; Guan, J. Transplanting Path Planning and Motion Functions Research of the High-Speed Tray Seedling Transplanting Robot. *Adv. Mater. Res.* **2013**, *694*, 1747–1752. [CrossRef]
88. Hu, J.; Yan, X.; Ma, J.; Qi, C.; Francis, K.; Mao, H. Dimensional Synthesis and Kinematics Simulation of a High-Speed Plug Seedling Transplanting Robot. *Comput. Electron. Agric.* **2014**, *107*, 64–72. [CrossRef]
89. Hu, J.; Jin, H.; Chang, Y.; Liu, W.; Han, L.; Yang, Q. Dimensional Synthesis and Trajectory Planning of Plug Seedling Transplanting Robot Based on Delta Parallel Mechanism. *Trans. Chin. Soc. Agric. Mach.* **2017**, *48*, 28–35. [CrossRef]
90. CMP Machine Videos CMP: Automated Seedling Planting System. Available online: https://www.youtube.com/watch?v=gbf84QfSKuA (accessed on 2 March 2023).
91. Hu, Y.; Wang, Z.; Zhang, H. Design and Test of an Automatic Seedling Transplanting Machine. *J. Agric. Mech. Res.* **2020**, *42*, 59–63+69. [CrossRef]
92. Tong, J.; Li, J.B.; Jiang, H.Y. Machine Vision Techniques for the Evaluation of Seedling Quality Based on Leaf Area. *Biosyst. Eng.* **2013**, *115*, 369–379. [CrossRef]
93. Ling, P.P.; Ruzhitsky, V.N. Machine Vision Techniques for Measuring the Canopy of Tomato Seedling. *J. Agric. Eng. Res.* **1996**, *65*, 85–95. [CrossRef]
94. Feng, Q.; Liu, X.; Jiang, K.; Fan, P.; Wang, X. Development and Experiment on System for Tray-Seedling on-Line Measurement Based on Line Structured-Light Vision. *Trans. Chin. Soc. Agric. Eng.* **2013**, *29*, 143–149.
95. Tai, Y.W.; Ling, P.P.; Ting, K.C. Machine Vision Assisted Robotic Seedling Transplanting. *Trans. ASAE* **1994**, *37*, 661–667. [CrossRef]

96. Jin, X.; Li, R.; Tang, Q.; Wu, J.; Jiang, L.; Wu, C. Low-Damage Transplanting Method for Leafy Vegetable Seedlings Based on Machine Vision. *Biosyst. Eng.* **2022**, *220*, 159–171. [CrossRef]
97. Syed, T.N.; Jizhan, L.; Xin, Z.; Shengyi, Z.; Yan, Y.; Mohamed, S.H.A.; Lakhiar, I.A. Seedling-Lump Integrated Non-Destructive Monitoring for Automatic Transplanting with Intel RealSense Depth Camera. *Artif. Intell. Agric.* **2019**, *3*, 18–32. [CrossRef]
98. Sadeghi-Tehran, P.; Sabermanesh, K.; Virlet, N.; Hawkesford, M.J. Automated Method to Determine Two Critical Growth Stages of Wheat: Heading and Flowering. *Front. Plant Sci.* **2017**, *8*, 252. [CrossRef]
99. Weng, Y.; Zeng, R.; Wu, C.; Wang, M.; Wang, X.; Liu, Y. A Survey on Deep-Learning-Based Plant Phenotype Research in Agriculture. *Sci. Sin.* **2019**, *49*, 698–716. [CrossRef]
100. Zhao, K.; Zhao, L.; Zhao, Y.; Deng, H. Study on Lightweight Model of Maize Seedling Object Detection Based on YOLOv7. *Appl. Sci.* **2023**, *13*, 7731. [CrossRef]
101. Samiei, S.; Rasti, P.; Ly Vu, J.; Buitink, J.; Rousseau, D. Deep Learning-Based Detection of Seedling Development. *Plant Methods* **2020**, *16*, 103. [CrossRef]
102. Wang, Y.; Yu, H.; Liu, Y. Mechanical Transplanting Plug Tray Localization Method Based on Binocular Stereo Vision. *Trans. Chin. Soc. Agric. Eng.* **2016**, *32*, 43–49. [CrossRef]
103. Yang, S.; Zheng, L.; Gao, W.; Wang, B.; Hao, X.; Mi, J.; Wang, M. An Efficient Processing Approach for Colored Point Cloud-Based High-Throughput Seedling Phenotyping. *Remote Sens.* **2020**, *12*, 1540. [CrossRef]
104. Xu, S.; Zhang, Y.; Dong, W.; Bie, Z.; Peng, C.; Huang, Y. Early Identification and Localization Algorithm for Weak Seedlings Based on Phenotype Detection and Machine Learning. *Agriculture* **2023**, *13*, 212. [CrossRef]
105. Yang, Z.; Zhang, W.; Li, W.; Chen, Y.; Song, P. Information Acquisition Method of Potted-Seedling Transplanting Fitness Using Monocular Vision. *Trans. Chin. Soc. Agric. Eng.* **2014**, *30*, 112–119. [CrossRef]
106. Yang, Z.; Zhang, W.; Li, W.; Chen, Y. Monocular Vision-Based Method for Direction Adjustment of Transplanting Potted-Seedling Leaves. *Trans. Chin. Soc. Agric. Eng.* **2014**, *30*, 26–33. [CrossRef]
107. Suzuki, T.; Murase, H. Non-Destructive Growth Measurement by Machine Vision for Cabbage Plug Seedlings Population with Missing Plants. *IFAC Proc. Vol.* **2000**, *33*, 151–156. [CrossRef]
108. Wang, Y.; Xiao, X.; Liang, X.; Wang, J.; Wu, C.; Xu, J. Plug Hole Positioning and Seedling Shortage Detecting System on Automatic Seedling Supplementing Test-Bed for Vegetable Plug Seedlings. *Trans. Chin. Soc. Agric. Eng.* **2018**, *34*, 35–41.
109. Jiang, H.; Shi, J.; Ren, Y.; Ying, Y. Application of Machine Vision on Automatic Seedling Transplanting. *Trans. Chin. Soc. Agric. Eng.* **2009**, *25*, 127–131.
110. Wen, Y.; Zhang, L.; Huang, X.; Yuan, T.; Zhang, J.; Tan, Y.; Feng, Z. Design of and Experiment with Seedling Selection System for Automatic Transplanter for Vegetable Plug Seedlings. *Agronomy* **2021**, *11*, 2031. [CrossRef]
111. Liu, J.; Liu, W.; Mao, H.; Li, P. Analysis on Picking Plant Sequence and Route of Transplanting Robotic for Column Cultivation. *Trans. Chin. Soc. Agric. Eng.* **2014**, *30*, 28–35. [CrossRef]
112. Jiang, Z.; Zhou, M.; Tong, J.; Jiang, H.; Yang, Y.; Wang, A.; You, Z. Comparing an Ant Colony Algorithm with a Genetic Algorithm for Replugging Tour Planning of Seedling Transplanter. *Comput. Electron. Agric.* **2015**, *113*, 225–233. [CrossRef]
113. Tong, J.; Jiang, H.; Zhou, M. Optimization of Seedling Transplanting Path Based on Genetic Algorithm. *Trans. Chin. Soc. Agric. Mach.* **2013**, *44*, 45–49+26. [CrossRef]
114. He, Y.; Yang, T.; Wu, C.; Yu, Y.; Tong, J.; Chen, C. Optimization of Replugging Tour Planning Based on Greedy Genetic Algorithm. *Trans. Chin. Soc. Agric. Mach.* **2017**, *48*, 36–43. [CrossRef]

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.

MDPI AG
Grosspeteranlage 5
4052 Basel
Switzerland
Tel.: +41 61 683 77 34

Agriculture Editorial Office
E-mail: agriculture@mdpi.com
www.mdpi.com/journal/agriculture



Disclaimer/Publisher's Note: The title and front matter of this reprint are at the discretion of the Guest Editors. The publisher is not responsible for their content or any associated concerns. The statements, opinions and data contained in all individual articles are solely those of the individual Editors and contributors and not of MDPI. MDPI disclaims responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.





Academic Open
Access Publishing

mdpi.com

ISBN 978-3-7258-3212-5