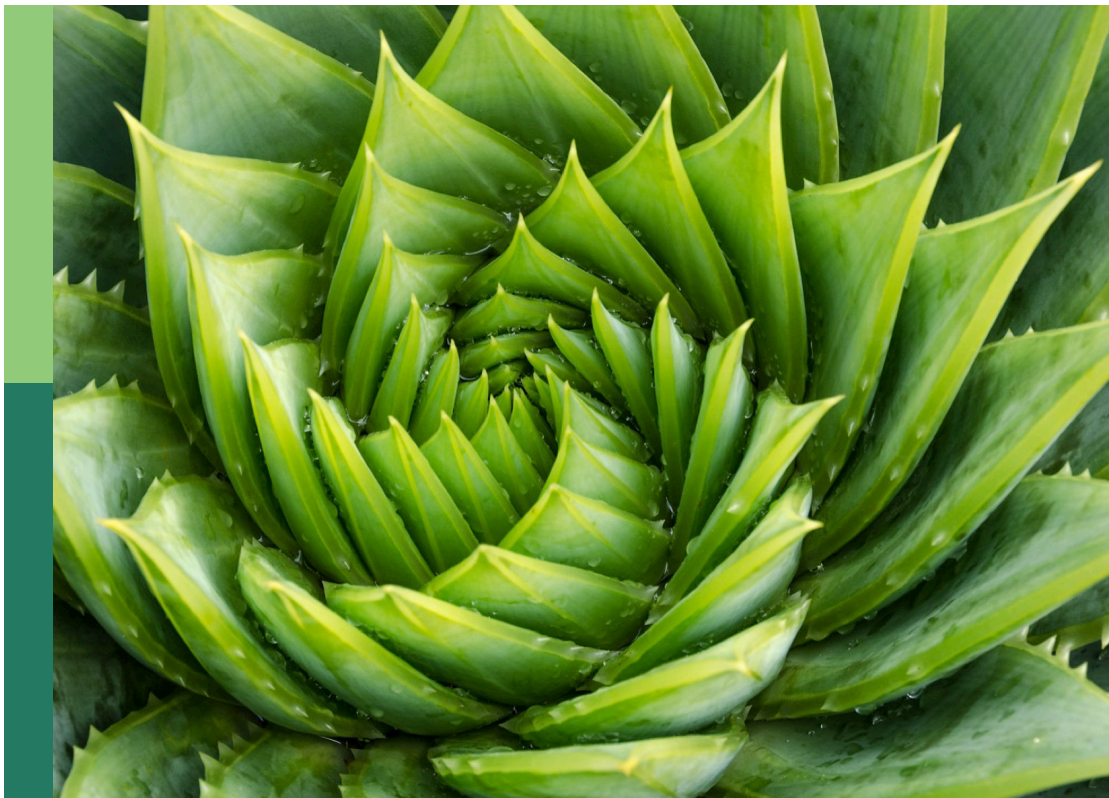
frontiers | Research Topics

Machine learning and artificial intelligence for smart agriculture

Edited by
Chuanlei Zhang, Dong Sun Park, Sook Yoon and Shanwen Zhang

Published in
Frontiers in Plant Science



ISSN 1664-8714
ISBN 978-2-83251-410-8
DOI 10.3389/978-2-83251-410-8

About Frontiers

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

Frontiers journal series

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

Dedication to quality

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

What are Frontiers Research Topics?

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

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

Machine learning and artificial intelligence for smart agriculture

Topic editors

Chuanlei Zhang — Tianjin University of Science and Technology, China
Dong Sun Park — Jeonbuk National University, Republic of Korea
Sook Yoon — Mokpo National University, Republic of Korea
Shanwen Zhang — Xijing University, China

Citation

Zhang, C., Park, D. S., Yoon, S., Zhang, S., eds. (2023). *Myeloid derived suppressor cells as disease modulators*. Lausanne: Frontiers Media SA.
doi: 10.3389/978-2-83251-410-8

Table of contents

frontiers | Frontiers in Plant Science

TYPE Editorial
PUBLISHED 09 January 2023
DOI 10.3389/fpls.2022.1121468

OPEN ACCESS

EDITED AND REVIEWED BY
Lei Shu,
Nanjing Agricultural University, China

*CORRESPONDENCE
Chuanlei Zhang
97313114@tust.edu.cn

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

RECEIVED 11 December 2022
ACCEPTED 20 December 2022
PUBLISHED 09 January 2023

CITATION
Zhang C, Park DS, Yoon S and Zhang S (2023) Editorial: Machine learning and artificial intelligence for smart agriculture.
Front. Plant Sci. 13:1121468.
doi: 10.3389/fpls.2022.1121468

Editorial: Machine learning and artificial intelligence for smart agriculture

Chuanlei Zhang[1]*, Dong Sun Park[2], Sook Yoon[3] and Shanwen Zhang[4]

[1]Artificial Intelligence College, Tianjin University of Science and Technology, Tianjin, China, [2]Department of Electronics Engineering, Jeonbuk National University, Jeonju, Republic of Korea, [3]Department of Computer Engineering, Muan, Republic of Korea, [4]Information Engineering College, Xijing University, Xi'an, China

KEYWORDS

computer vision, internet of things, agricultural robot, pest control, food security

Editorial on the Research Topic
Machine learning and artificial intelligence for smart agriculture

1 Introduction

Information, knowledge and equipment are the three main components of smart agriculture. This special edition focuses on a few issues that still require research and discussion. For example, using and enhancing machine learning techniques for crop disease and pest detection and recognition, plant species recognition, smart agricultural IoT, food material supply chain security tracing, and other crucial issues in smart agriculture.

2 Computer vision

Visual recognition technology has been increasingly applied to numerous areas of agricultural development with the advancement of computer graphics and image processing technology in artificial intelligence (AI). Today, there is still a significant room for this technology in modern agriculture (Tombe, 2020; Benos et al., 2021; Dhanya et al., 2022).

During the process of green apple harvesting or yield estimation, the accurate recognition and fast location of the target fruit bring tremendous challenges to the vision system. Sun et al. improved a density peak cluster segmentation algorithm for RGB images with the help of a gradient field of depth images to locate and recognize target fruit. Specifically, the image depth information is adopted to analyze the gradient field of the target image.

There are two serious apple diseases which are ring rot produced by Botryosphaeria dothidea and anthracnose caused by Colletotrichum gloeosporioides. In Feng et al., the

automatic separation between two diseases was examined using image processing technologies. The acquired disease images were preprocessed using morphological opening and closing reconstruction, color image contrast stretching, and image scaling. Then, two crop leaf lesion segmentation algorithms based on circle fitting were suggested and applied. Support vector machine (SVM) models and random forest models were used based on individual LBP histogram features and various LBP histogram feature combinations.

In addition, crop classification is crucial for the development of phenotypes and genetic resources. Fu et al. created rapeseed dataset (RSDS) using eight categories of data gathered. The target-dependent neural architecture search (TD-NAS) was proposed. TD-NAS is a revolutionary target-dependent search technique based on VGGNet.

3 Internet of things

Sensor nodes in agricultural IoTs, such as soil temperature and humidity sensors, air temperature and humidity sensors, etc., collect data in the agricultural environment. The data is wireless transferred from the sensor nodes to the sink node for data collection. The gateway change the protocol into one that can be communicated over the Internet when it receives data from the sink node (Ayaz et al., 2019; Doshi et al., 2019; Priya et al., 2021).

Usually, multidimensional time series data are produced in enormous quantities by smart agricultural IoT. However, due to the limitations of the scenarios, data loss and misrepresentation are frequent problems with smart agricultural IoTs. In order to solve the aforementioned issues, in, using generative adversarial networks (GAN), Cheng et al. offered a new anomaly detection model that can handle the multidimensional time series data produced by smart agricultural IoTs.

Meanwhile, a multi-objective strategy based on supervised machine learning was utilized in Uyeh et al. to identify the ideal number of sensors and installation locations in a protected cultivation system. A tree-based model in the form of a gradient boosting technique was specifically adapted to observed (temperature and humidity) and derived circumstances (dew point temperature, humidity ratio, enthalpy, and specific volume). Time series forecasting was used for feature variables. In order to choose the right number and locations for the best sensors in a protected cultivation system, a machine learning model was created and put forth.

In larger fields, such as those seen in large-area agriculture, more sensors/nodes are advised to better account for soil heterogeneity. But farmers must pay a higher and frequently prohibitive price for this (purchase, labor costs from installation and removal, and maintenance). The agricultural industry would benefit greatly from methodologies that allow for maintaining monitoring capability/intensity with fewer in-field sensors. In Maia et al., sensor data analysis over two irrigation seasons in three cotton fields from two cotton-growing regions of Australia revealed a connection between soil matric potential and cumulative crop evapotranspiration (ETcn) derived from satellite measurements between irrigation events. This connection can be represented as a second-degree function.

4 Agricultural robots

Globally, the use of intelligent machines and robots in agriculture is receiving increasing attention (Oliveira et al., 2020). A picking robot is a type of agricultural robot that uses a variety of sensors to sense the complicated agricultural environment and then picks the target using this knowledge and a decision-making algorithm. Ma et al. explore the distributed averaging issues of agriculture picking multi-robot systems under directed communication topologies by utilizing the sampled data. A distributed protocol based on nearest-neighbor information is presented using the principles of algebraic graph theory and matrix theory.

5 Pest control

Crop diseases and pests have long been a significant issue in agricultural production, having a negative impact on farmers' income, modern agricultural development, and agricultural output. Early disease and pest identification, monitoring, and control are crucial for preventing the large-scale spread of diseases and pests, preserving the quality of crops, and reducing environmental pollution brought on by pesticide residues (Buja et al., 2021; Liu and Wang, 2021).

The brown planthopper (BPH), Nilaparvata lugens (Stl; Hemiptera: Delphacidae), is a piercing-sucking insect that seriously harms rice plants by sucking out their phloem sap and spreading viruses. For reducing mating rates, a physical control mechanism based on BPH courting disruption is a viable approach to reducing environmental pollution. To gather effective courtship disrupting signals. Feng et al. created a vibration signal recording, monitoring, and playback system for BPHs. This technology was used to gather and evaluate male competitiveness and BPH courting signals in order to determine their frequency spectra. According to the findings, the mean main vibration frequency and mean pulse rate of female courtship signals are 234 Hz and 23 Hz, respectively. Male courting signals had mean main vibration and pulse frequencies of 255 Hz and 82 Hz, respectively.

Furthermore, Cnaphalocrocis medinalis, Sogatella furcifera, and Nilaparvata lugens are three kinds of migratory pests that severely reduce rice yield and result in economic losses each year. Sun et al. create an intelligent monitoring system of migrating pests based on searchlight trap and machine vision to replace manual identification of migratory pests in. The system consists

of a cloud server, a Web client, a migratory pest automatic identification model, and a searchlight trap based on machine vision. The searchlight trap uses lights at night to draw in high-altitude migrating insects. All captured insects are distributed using rotary brushes and multi-layer insect conveyor belts. The intelligent monitoring system can automatically monitor the three migratory pests in time.

6 Food security

For the quality and safety of agricultural products, food traceability is crucial (Ivar et al., 2020). In Jing and Li a hybrid mode of block-chain and IoTs is used to build a traceability system for red jujube. The solution addresses the issue of date and quality traceability by integrating block-chain and IoTs technologies with properties of tamper-proof, decentralization, and distributed storage. The entire process of red jujube cultivation, processing, and sales is documented in the block. To guarantee the realization of quality traceability of red jujube in the framework throughout the whole process of big data processing, and the crucial data gathered in each procedure is saved in the database.

Author contributions

The four professors contributed to edit and review the manuscripts for this edition. CZ was mainly responsible for the computer vision and food security part, DP was responsible for the Iot and robot, SY was responsible for the pest control, and SZ was responsible for the Iot and computer vision part. All authors contributed to the article and approved the submitted version.

Funding

This work is supported by the National Natural Science Foundation of China (Nos. 62172338 and 62273284).

Conflict of interest

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

Publisher's note

References

Ayaz, M., Ammad-Uddin, M., Sharif, Z., Mansour, A., and Aggoune, E. -H. M.. (2019). Internet-of-Things (IoT)-based smart agriculture: Toward making the fields talk. *IEEE Access* 7, 129551–129583. doi: 10.1109/ACCESS.2019.2932609

Benos, L., Tagarakis, A. C., Dolias, G., Berruto, R., Kateris, D., and Bochtis, D. (2021). Machine learning in agriculture: A comprehensive updated review. *Sensors* 21, 3758. doi: 10.3390/s21113758

Buja, I., Sabella, E., Monteduro, A. G., Chiriacò, M. S., De Bellis, L., Luvisi, A., et al. (2021). Advances in plant disease detection and monitoring: From traditional assays to in-field diagnostics. *Sensors* 21, 2129. doi: 10.3390/s21062129

Dhanya, V. G., Subeesh, A., Kushwaha, N. L., Vishwakarma, D. K., NageshKumar, T., Ritika, G., et al. (2022). Deep learning based computer vision approaches for smart agricultural applications. *Artif. Intell. Agric.* 6, 211–229. doi: 10.1016/j.aiia.2022.09.007

Doshi, J., Patel, T., and Bharti, S. K. (2019). Smart farming using IoT, a solution for optimally monitoring farming conditions. *Proc. Comput. Sci.* 160, 746–751. doi: 10.1016/j.procs.2019.11.016

Ivar, V., Shah, A. N., and Sofia, B. (2020). Food security, safety, and sustainability–getting the trade-offs right. *Front. Sustain. Food Syst.* 4. doi: 10.3389/fsufs.2020.00016

Liu, J., and Wang, X. (2021). Plant diseases and pests detection based on deep learning: A review. *Plant Methods* 17, 22. doi: 10.1186/s13007-021-00722-9

Oliveira, L. F., Silva, M., and Moreira, A. (2020). Agricultural robotics: A state of the art survey. 23rd international conference series on climbing and walking robots and the support technologies for mobile MachinesAt: Moscow, Russian federation 279–286. doi: 10.13180/clawar.2020.24-26.08.44

Priya, O.V., and Ramanujam, S. (2021). Impact of Internet of things (IoT) in smart agriculture. *Recent Trends in Intensive Computing*. doi: 10.3233/APC210176

Tombe, R. (2020). Computer Vision for Smart Farming and Sustainable Agriculture. *Proceedings of the 2020 IST-Africa Conference (ISTAfrica)*. 2020 May 18–22; IEEE: Piscataway, NJ, USA.

frontiers | Frontiers in Plant Science

ORIGINAL RESEARCH
published: 06 June 2022
doi: 10.3389/fpls.2022.890563

Anomaly Detection for Internet of Things Time Series Data Using Generative Adversarial Networks With Attention Mechanism in Smart Agriculture

Weijun Cheng[1], Tengfei Ma[1], Xiaoting Wang[1] and Gang Wang[2]*

[1]School of Information Engineering, Minzu University of China, Beijing, China, [2]Key Laboratory of Mining Disaster Prevention and Control, Shandong University of Science and Technology, Qingdao, China

OPEN ACCESS

Edited by:
Shanwen Zhang,
Xijing University, China

Reviewed by:
Le Li,
Cornell University, United States
Fuhong Lin,
University of Science and Technology Beijing, China
Ashesh Gaur,
University of Oklahoma, United States

***Correspondence:**
Weijun Cheng
weijuncheng@muc.edu.cn

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

Received: *06 March 2022*
Accepted: *17 May 2022*
Published: *06 June 2022*

Citation:
Cheng W, Ma T, Wang X and Wang G (2022) Anomaly Detection for Internet of Things Time Series Data Using Generative Adversarial Networks With Attention Mechanism in Smart Agriculture. Front. Plant Sci. 13:890563. doi: 10.3389/fpls.2022.890563

More recently, smart agriculture has received widespread attention, which is a deep combination of modern agriculture and the Internet of Things (IoT) technology. To achieve the aim of scientific cultivation and precise control, the agricultural environments are monitored in real time by using various types of sensors. As a result, smart agricultural IoT generated a large amount of multidimensional time series data. However, due to the limitation of applied scenarios, smart agricultural IoT often suffers from data loss and misrepresentation. Moreover, some intelligent decision-makings for agricultural management also require the detailed analysis of data. To address the above problems, this article proposes a new anomaly detection model based on generative adversarial networks (GAN), which can process the multidimensional time series data generated by smart agricultural IoT. GAN is a deep learning model to learn the distribution patterns of normal data and capture the temporal dependence of time series and the potential correlations between features through learning. For the problem of generator inversion, an encoder–decoder structure incorporating the attention mechanism is designed to improve the performance of the model in learning normal data. In addition, we also present a new reconstruction error calculation method that measures the error in terms of both point-wise difference and curve similarity to improve the detection effect. Finally, based on three smart agriculture-related datasets, experimental results show that our proposed model can accurately achieve anomaly detection. The experimental precision, recall, and F1 score exceeded the counterpart models by reaching 0.9351, 0.9625, and 0.9482, respectively.

Keywords: anomaly detection, smart agriculture, time series data, deep learning, generative adversarial network, attention mechanism

INTRODUCTION

Nowadays, Internet of Things (IoT) technology has been obtained rapidly developments, as a paradigm, to drive the evolution of modern industries and smart cities. As for serious challenges in environmental pollution, energy depletion, and water shortage in the whole world, there is an urgent need for the agriculture industry to move toward digitalization (Cao et al., 2021). To address these challenges, smart agriculture solutions based on real-time monitoring and decision-making have been received increasing attention. Smart agriculture is a deep combination of IoT technology and modern agriculture, which mainly takes modern agriculture as an application scenario and applies IoT technology to achieve a goal of scientific cultivation and precise control (Farooq et al., 2019).

For smart agriculture IoT systems, automated management and smart decision of IoT applications are driven by the detailed analysis of data (Cao et al., 2021). These data are collected by a large number of various types of sensors and provide information about different environmental conditions. Thus, environmental monitoring and data analysis play an important role in increasing crop yields. The sensors in different application scenarios are shown in **Figure 1**. However, IoT devices in smart agriculture are usually exposed to harsh environments and are highly susceptible to damage due to cost control (Rafii and Kechadi, 2019; Abdallah et al., 2021). In addition, the heterogeneous nature of network devices makes it difficult to design protocols, and the transmission of data is easily compromised (Pundir and Sandhu, 2021). Poor communication quality can lead to data loss and misrepresentation. Increasingly complex IoT systems bring technical complexity and therefore make the design of privacy and security mechanisms more difficult. This can also expose the network to attacks that could lead to data tampering (Abdallah et al., 2021). Missing or misrepresented data is significantly different from normal data in the time series data collected by the sensors (Moso et al., 2021). These can be considered as anomalies in the data (Adkisson et al., 2021). Moreover, IoT applications also require an algorithm to analyze these data to facilitate intelligent decision-making. By analyzing the data in detail, the intelligent system can make the most efficient resource scheduling to increase crop yield. Testing for crop growth patterns can help reduce soil depletion, and different weather and soil conditions can affect irrigation decisions (Vilenski et al., 2019; Vyas and Bandyopadhyay, 2020; Garg et al., 2021). The main idea of data analysis in the smart agriculture scenario is to analyze various sensor data, and the analysis results can reflect the changes in the environment (Khalil et al., 2021). In particular, data that differ from normal data due to environmental changes can be also designated as anomalies. Therefore, anomaly detection has become an important work of smart agricultural IoT.

The data collected by smart agriculture IoT sensors is mainly called stream data, also called time series data. They are a series of infinite data points with a timestamp T . The purpose of time series anomaly detection is to find anomaly points or anomaly subsequences in a time series. In previous years, machine learning-based data mining techniques have been evaluated and achieved high performance in anomaly detection (Nassif et al., 2021; Pang et al., 2021). Due to the specificity of time series, their data sets lack labeling information. Therefore, it is usually treated as unsupervised machine learning. Unsupervised tasks do not require expert knowledge and can automatically adapt to data changes (Yu et al., 2021). Different environmental indicators generate different time series data. Multiple time series form multidimensional time series data, which is the main data format in smart agriculture. There are potential correlations between the different dimensions of these data, which should be considered primarily. However, existing unsupervised machine learning methods cannot handle the non-linearity of potential correlations in multidimensional time

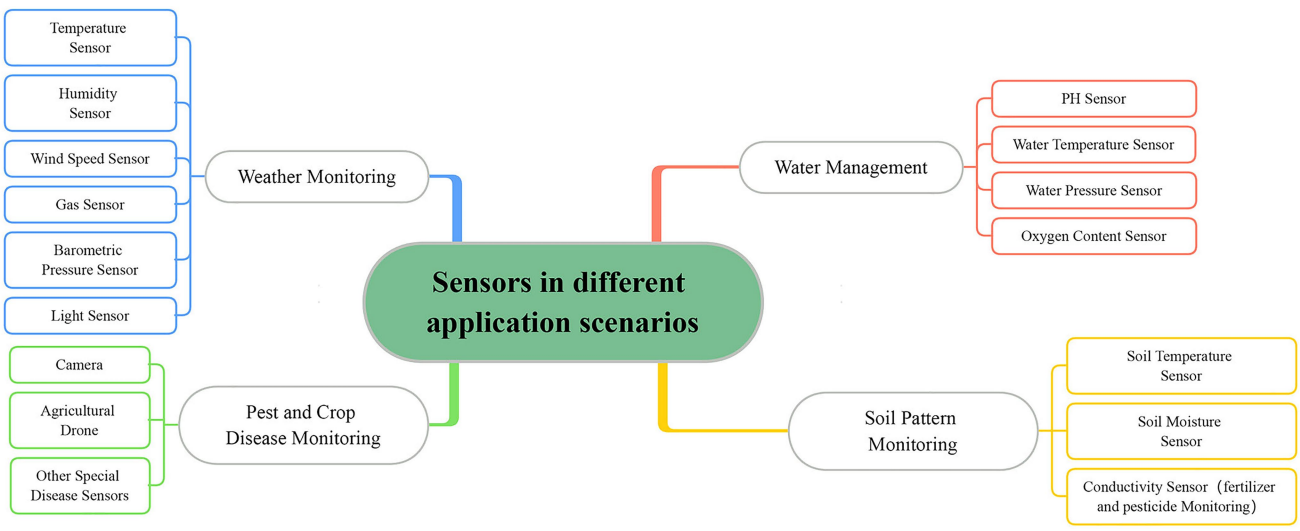


FIGURE 1 | Sensors in different application scenarios.

series (Li et al., 2019). Up to date, with the increasing number and types of sensors, smart agriculture generated a large amount of time series data, and there exist two challenges in anomaly detection. One is that the amount of monitored variables and data points is exploding, and the other is that there are potential correlations and time dependencies between multidimensional variables. They lead to high-dimensional and heterogeneous time series data features, which cannot be accomplished by machine learning-based anomaly detection models (Yang et al., 2021). Thus, it is necessary to find some new research approaches to solve these emerging problems of anomaly detection in smart agriculture.

In recent years, deep learning has been proposed for anomaly detection, and most of them are reconstruction-based models. The general process of reconstruction-based anomaly detection is that a model is employed to learn the distribution of normal data, and then the trained model is used to reconstruct the data to be measured. The error between the reconstructed data and the original data is used to determine if the data is anomalous. Reconstruction-based anomaly detection models can model large-scale data and capture potential correlations between multidimensional data. Among them, generative adversarial networks (GAN) work well (Goodfellow et al., 2014). GAN generally contain generator and discriminator. The generator can generate samples, and the discriminator can determine whether the input sample is the original sample or the sample generated by the generator. The generator wants to generate samples that are closest to the original samples to fool the discriminator. The discriminator wants to accurately determine whether the sample is a real sample or not. The learning ability of the model is continuously improved by the adversarial learning of both. GAN was initially introduced to anomaly detection to solve problems related to image data (Schlegl et al., 2017; Zenati et al., 2018a). With the growth of the number and dimensionality of time series data, GAN was introduced to time series data anomaly detection due to its superiority in processing high-dimensional data. GAN-based anomaly detection belongs to the reconstruction-based anomaly detection models, in which GAN is used to learn the distribution of normal data. A trained GAN to reconstruct anomaly data will produce large reconstruction errors. Finally, the anomaly score is used to determine whether the test sample is anomalous, where the anomaly score mainly includes the reconstruction error.

Li et al. (2019) and Bashar and Nayak (2020) introduced the general GAN into the anomaly detection model for time series data. The goal of these algorithms is to detect time series data anomalies quickly and accurately by GAN. Generally, there are two main types of anomalies: one is data loss or data misrepresentation caused by equipment failure or network anomalies, and the other is data anomalies that do not conform to the potential correlation of normal data distribution. However, since the generator input of GAN is random normal data, this brings inconvenience to the calculation of reconstruction error. The calculation of each reconstruction error requires finding the optimal normal data corresponding to the reconstructed samples, which needs the inversion of the generator. This leads to a large computational cost and may also degrade the detection results. Some works in the field of anomaly detection have focused more on changes in model structure, but there have been few improvements to the way errors are calculated. Most studies considered only a single computational method, and the point-wise difference calculation was widely adopted (Li et al., 2019; Bashar and Nayak, 2020; Geiger et al., 2020). This does not exactly fit the time series data format and sometimes does not conform to the true definition of error. Time series data is a series of data points that can form a smooth curve. For the curve as a form of data, the curve similarity should be considered as an error measure.

Motivated by the above observation, we focus on the anomaly detection of multidimensional time series data, which is generated from different sensor data in smart agricultural systems. In this paper, we propose a new GAN-based anomaly detection method. In particular, for generator inversion, an encoder–decoder architecture is designed. In this architecture, we introduce an attention mechanism that can effectively improve the reconstruction effect. Then, a new reconstruction error calculation is provided. The point-wise difference and curve similarity are jointly considered as reconstruction errors, which makes the error definition more realistic and improves detection performance. Finally, we conduct experiments using three data sets related to smart agriculture and specialize the model parameters according to the data set characteristics. The experimental results show that our approach outperforms the other four counterpart anomaly detection methods.

RELATED WORK

With the rapid development of computer technology, researchers began to experiment with computer technology to solve anomaly detection. Hawkins (1980) had a widely accepted explanation of anomalies, namely, "in a given data set, anomalous data are that part of the data that is significantly different from the majority of the data." Current anomaly detection methods can be broadly classified into proximity-based methods, probability-based methods, prediction-based methods, and deep learning-based methods (Aggarwal, 2017). These methods except deep learning are called traditional methods. They used statistical measures to calculate the correlation between the data records. These techniques assumed that the time series is linear and follows a known statistical distribution, which makes them inapplicable to many practical problems (Adhikari and Agrawal, 2013). As the volume and dimensionality of data grow, more deep learning algorithms have been proposed for anomaly detection on complex data. Deep learning-based anomaly detection methods have advantages over these methods in characterizing multidimensional time series data and are more helpful in solving practical problems.

Generally, time series data anomaly detection algorithms are divided into two steps. The data are modeled by different data structures and then the degree of deviation of the test data from the normal data is evaluated based on different forms of metrics (e.g., distance-based and density-based). Data with excessive deviations are judged to be abnormal. The deep

learning-based anomaly detection is similar to the above process. Neural networks are used in the data representation phase to learn the data distribution, and reconstruction-based methods are applied in the anomaly calculation phase to measure error. Since the data structure learns the normal data distribution, there will be a large reconstruction error using this model to reconstruct abnormal data. In recent years, there has been an increasing number of studies using GAN for anomaly detection. As a result, a state-of-the-art survey of the anomaly detection for GAN is discussed in the following section.

GAN for Image Anomaly Detection

AnoGAN (Schlegl et al., 2017) was the first work that applied GAN for image anomaly detection. The model was trained with normal data and the final anomaly scores were obtained by calculating errors in the trained generator and discriminator. The reconstruction error was calculated in the generator to calculate the error more efficiently. However, this computation requires finding the inverse mapping from the data space to the latent space and is not synchronized with the training. It can lead to extremely high error computation time. Zenati et al. (2018a) proposed an efficient GAN-based anomaly detection to solve the above problem. They added an extra encoder to GAN to avoid looking for latent vector at each detection. The calculation of the anomaly error was the same as the AnoGAN. Skip-GANomaly proposed by Akçay et al. (2019) introduced an architecture of skip connection to improve image reconstruction. The model improved image reconstruction but did not perform well on all data sets, which was limited by unstable training. Zenati et al. (2018b) pointed out that AnoGAN was inappropriate for real-time anomaly detection or larger data sets. They proposed a bi-directional GAN for image anomaly detection, which simultaneously trained the inverse mapping through an encoder network. The model contained three discriminators which effectively improved training stability.

GAN for Time Series Anomaly Detection

The achievements of GAN in image anomaly detection have attracted the attention of researchers, and have been introduced into time series anomaly detection. Li et al. (2018) proposed a GAN-based anomaly detection method (GAN-AD) for time series data, which was used to detect possible anomalous behaviors in complex networks. To capture the correlation of time series data, Long Short-Term Memory networks (LSTM) were used as the basic model to learn normal data distribution patterns. For the evaluation of the error, since the output of the discriminator indicated whether the sample is false or not, it was used directly as the anomaly score to find the anomaly. Li et al. (2019) later extended their study to use a vanilla GAN model to capture multivariate time series model distributions and detect anomalies using reconstruction errors and discriminator outputs. Bashar and Nayak (2020) improved on AnoGAN (Schlegl et al., 2017) and proposed an anomaly detection algorithm for time series data. The model used a convolutional neural network (CNN) as the basic network to capture the correlation between variables. Both of the above models can learn the time correlation of time series data and effectively detect anomalies. However, they also need to find the inverse mapping from the real space to the latent space, which requires an inversion of the generator resulting in a longer computation time.

To address this problem, Geiger et al. (2020) proposed TadGAN based on Zenati et al. (2018a). This model introduced cycle-consistent GAN architectures, which allowed the generator to compute the reconstruction error directly without finding the inverse mapping and reducing the computation time. For the calculation of the anomaly score, the combination of point-wise difference and discriminator was typically considered as the anomaly score. However, using point-wise difference measures alone does not exactly fit the time series data characteristics. The time series can form different smoothing curves, and the shape differences between these curves should be equally considered in the reconstruction error calculation. The TadGAN used curve similarity as a form of calculation of reconstruction error. However, they studied the point-wise difference and curve similarity separately and did not consider them together to meet a realistic definition. Apart from this, the existing articles are insufficient for the study of reconstruction errors.

Anomaly Detection in Smart Agriculture

Research on deep learning-based anomaly detection for IoT systems has yielded excellent results, some of which have been introduced into smart agriculture to address emerging challenges. Most of the research on smart agriculture has focused on the field of anomaly detection in agricultural images, such as the identification of pests and crop diseases. TPest-RCNN proposed by Li et al. (2021) aimed to identify whitefly and thrips in greenhouses. The model was trained on a set of pest images captured by a flytrap and used a transfer learning strategy to achieve improved detection. Liu and Wang (2020) optimized the feature layer of Yolo V3 model by using the image pyramid to achieve multi-scale feature detection and improved the detection accuracy and the speed of Yolo V3 model. Experiments showed that the model can accurately and quickly detect the location and category of tomato pests and diseases.

For time series data generated in agricultural IoT systems, some researchers have focused on anomaly detection of sensor network data. Several papers have offered specifics on anomalies in smart ecosystems (Cook et al., 2019; Hasan et al., 2019; Park et al., 2021). In smart agriculture scenarios, agricultural IoT devices are often exposed to harsh conditions that can lead to failure of the device itself, compromised communications, or malicious attacks, which can lead to data anomalies. Adkisson et al. (2021) proposed an anomaly detection model for smart farming using an unsupervised autoencoder machine learning model. The model used an autoencoder to encode and decode the data, and anomalous data generated a high reconstruction loss value. Ultimately, the test data was determined to be anomalous based on a threshold value. Abdallah et al. (2021) applied autoregressive integrated moving average (ARIMA) and LSTM model to a smart agricultural system and specialized models based on sensor constraints. The transfer learning strategies were introduced into the models to improve

the prediction. Anomaly detection for time series data can also be applied in crop harvesting. Moso et al. (2021) proposed a powerful ensemble-based approach for anomaly detection, which was mainly used for data streams generated in smart agriculture. This technology can be applied to crop data sets and identify anomalies that affect crop harvest.

In summary, time series anomaly detection based on deep learning has obtained excellent results, and GAN for anomaly detection has been continuously explored. Various deep learning models have been introduced to solve the problem of time series data for smart agriculture. As the amount and dimensionality of data increase, existing smart agriculture anomaly detection models are unable to handle the data. GAN has been introduced to various fields to process multidimensional time series data with better results. Therefore, we introduce GAN into smart agriculture for anomaly detection and specialize model structure for smart agriculture data characteristics. However, the existing studies of GAN for anomaly detection are limited by the problem of generator inversion and the reconstruction error is calculated in a simple way. To address the problem of GAN for anomaly detection, we design a new architecture and an error calculation method to improve the anomaly detection performance.

PROPOSED APPROACH

In this section, we first describe a novel GAN-based anomaly detection model and focus on how it uses an adversarial learning architecture by considering the dependencies between time series data. Then, the internal detailed architecture of the model is shown, which includes the encode–decoder in the generator and the structure of the discriminator. These designs have strong relevance to the goals of improving reconstruction effects and reducing error computation time. To better learn the data distribution, a multi-channel attention mechanism is embedded in the encoder and decoder, which can further improve the reconstruction effect. Finally, we introduce a new error calculation method in this model, which can describe the errors more rationally and improve the detection results effectively.

The core idea of the reconstruction-based anomaly detection method is to encode a data point (time series data in this model) and then decode the encoded data point to reconstruct the data. Anomalous data loses a lot of information during the encoding–decoding process, because what the model should learn is how to reconstruct normal data. Thus, a normal trained model cannot reconstruct abnormal data in the same way as normal data. Large reconstruction errors will arise in the process of reconstructing anomalous data. This means that the reconstructed data has a large difference from the original data. In this paper, GAN model is used to model the data in an attempt to learn the normal distribution of the data.

The basic task of anomaly detection is to identify whether the data to be tested conforms to the distribution of normal data, and data that do not conform to the normal distribution are defined as anomalous (Chalapathy and Chawla, 2019; Kwon et al., 2019). In this work, the completed trained GAN is used for anomaly detection. The test samples are processed in the same data processing manner and then fed into the model in an attempt to reconstruct them. The anomaly score is calculated using a jointly trained generator and discriminator, which consists of the output of the discriminator and the reconstruction error of the generator. For reconstruction errors, we use a new calculation to detect potential anomalies in the data (more details will be described in "Anomaly Detection").

The Proposed GAN Framework

The general architecture of our proposed model is shown in **Figure 2**. The first objective of this model is to learn the normal distribution of a given data set by means of adversarial training. Previous studies have taken random normal vectors in the latent space Z and inputted them into the generator for training (Li et al., 2019; Bashar and Nayak, 2020). The trained generator is able to implicitly capture the multivariate distribution of the training data and learn the mapping of random data to normal data. However, the error between the reconstructed time series data and the real data needs to be calculated in anomaly detection. In order to reconstruct the data, it is necessary to find the random normal data corresponding to the reconstructed data at each calculation. This process requires the inversion of the generator and is time-consuming and computationally resource intensive. To solve this problem, an encoder–decoder architecture is designed as a generator, in which the encoder learns the mapping of normal data to latent vector in the latent space and optimizes the computation time of reconstruction errors.

Sensors in smart agricultural systems perform continuous measurement tasks to detect changes in the environment. Therefore, they generate a large amount of multivariate time series data. We use LSTM as the basic model of generator and discriminator to deal with complex multidimensional time series data. For the characteristics of multidimensional time series data, the data streams are not processed separately. The entire data set is processed concurrently to capture potential interactions between variables. Multivariate time series data are divided into subsequences that are fed into the model through a sliding window mechanism. We set the window size as the super parameter of the model to determine the optimal window length, which can capture the data distribution in different situations according to the characteristics of different data sets. Here, the window size is set as

$$s_w = 30 \times i, i = 1, 2, \cdots 10.$$

The data first needs to be preprocessed before training. The multivariate time series data $\Phi \subset R^{T \times N}$ of the length T and number N of variables are partitioned into a training set $\Phi_{train} \subseteq R^{T_1 \times N}$, a validation set $\Phi_{validation} \subseteq R^{T_2 \times N}$ and a test set $\Phi_{test} \subseteq R^{T_3 \times N}$. Noted that the training set data must be all normal data. Next, the training data set $\Phi_{train} \subseteq R^{T_1 \times N}$ is divided into a series of subsequences $X_{train} = \left\{ X^i_{train}, \mathrm{i} = 1, 2, \cdots m \right\} \subseteq R^{S_w \times N}$ using a sliding window of size s_w, where X_{train} denotes $X_{t-s_w:t}$.

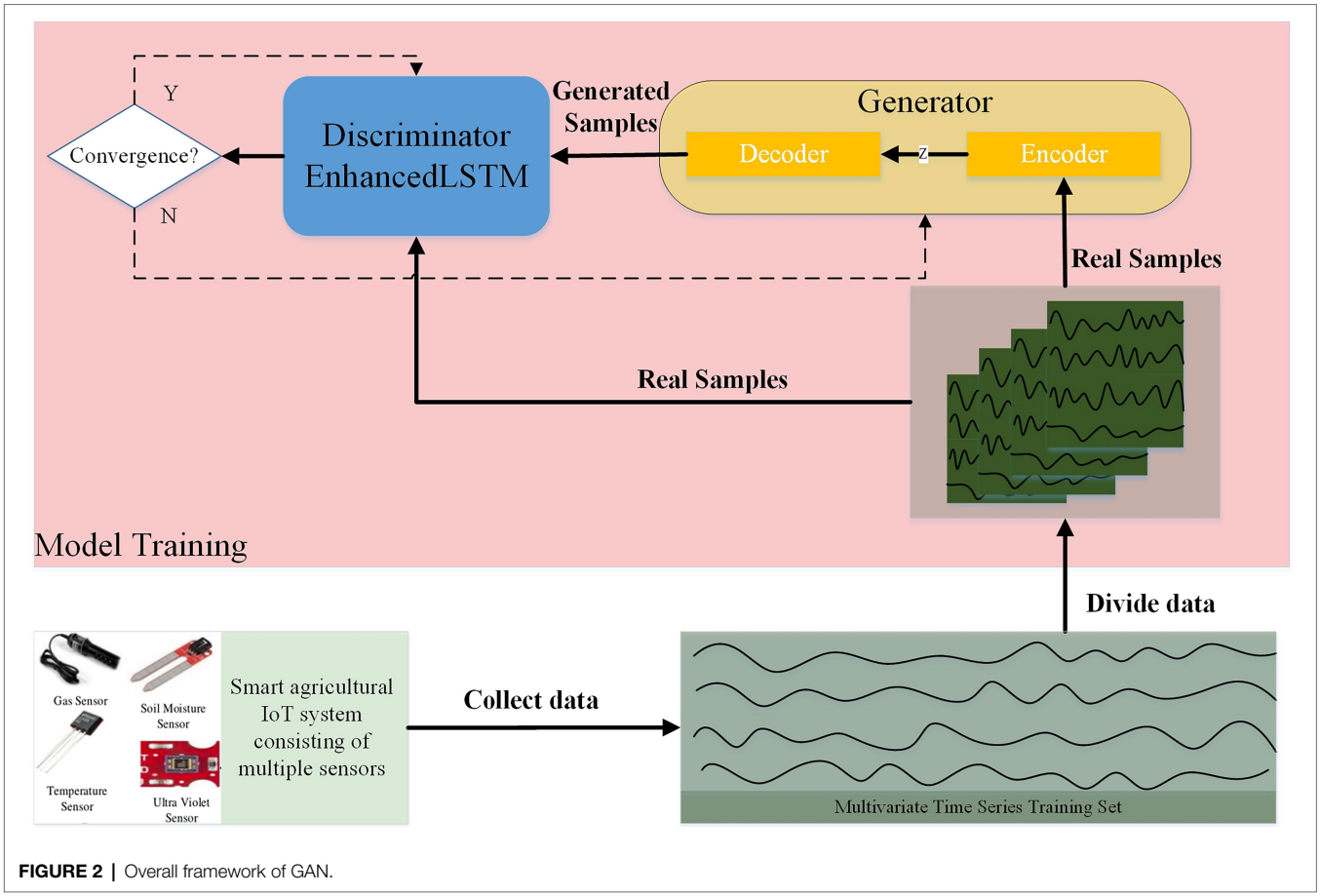


FIGURE 2 | Overall framework of GAN.

Given the step size s_t, the number of subsequences can be calculated by $m = \left[\left(T_1 - s_w \right) / s_t \right] + 1$. Similarly, the validation set $\Phi_{validation} \subseteq R^{T_2 \times N}$ can be partitioned into a series of subsequences $X_{validation} = \left\{ X^j_{validation}, j = 1,2,\cdots n \right\} \subseteq R^{S_w \times N}$, where $n = \left[\left(T_2 - s_w \right) / s_t \right] + 1$. Subsequences in the validation data set are marked to indicate whether the sequence is abnormal or not (1 means normal, 0 means abnormal). The test set is handled exactly in the same way as the validation set.

Model training is performed after the data preprocessing is completed. The distribution of the data is learned by the GAN model in adversarial training. In our model, the mapping functions of the X and Z domains are learned as $\varepsilon : X \to Z$ and $G : Z \to X$, respectively. X is the input data, which represents the training samples $\left\{ \left(x_i^{1 \cdots N} \right) \right\}_{i=1}^{t} \in X$ given by the model. Z is the vector in latent space and the encoder learns the mapping $\varepsilon : X \to Z$ to encode the input data as a latent vector. The mapping $G : Z \to X$ is learned by the decoder, which reconstructs the vector in latent space to the input data. With the above two mapping functions, we can achieve the data reconstruction: $x_i \to \varepsilon \left(x_i \right) \to G\left(\varepsilon \left(x_i \right) \right) \approx x_i$. These two mapping functions are obtained by adversarial learning methods, and together they form the generator of the GAN architecture.

The generator tries to deceive the discriminator by generating the real sample through the encoder–decoder architecture so that the discriminator judges the generated data as the real sample. To ensure that the distribution pattern of normal data is learned by the model, we make sure that the training data are all normal during the training phase. Unlike the general GAN that inputs the variables in latent space to the generator, the normal sample x_i after data segmentation processing is directly fed into the generator and the two mapping functions ε and G mentioned above learn the mapping patterns of the two stages, respectively. The samples are reconstructed as much as possible to the original samples after two mappings. Both the generator output $G\left(\varepsilon \left(x_i \right) \right)$ and the original data x_i are then sent to the discriminator to distinguish whether they are generated data or not. The generator tries to generate the same samples as the original data, and the discriminator tries to distinguish the real samples from the generated samples. This process is similar to the one in which the generator G uses the discriminator D as an adversary (Goodfellow et al., 2014). Adversarial training of both G and D continuously improves their performance until a set number of iterations is reached or the model converges. After continuous iterations of adversarial learning, the generator implicitly learns the normal data distribution and the discriminator can distinguish the real data from the

generated data. The overall loss function in this process is mainly adversarial loss.

Adversarial Loss

Both the generator and the discriminator try to optimize the competing loss functions during training. Thus, the optimization process can be considered as a minimax game problem. During the game, the generator tries to minimize the loss to make the generated sample as close as possible to the original sample. The discriminator tries to maximize the loss to distinguish the real samples from the generated samples. The adversarial loss of the training process is defined as follows:

$$\boldsymbol{L_{adv}} = \boldsymbol{E_{X\sim pX}}\left[\log\left(\boldsymbol{D}(\boldsymbol{X})\right)\right] + \boldsymbol{E_{X\sim pX}}\left[\log\left(\boldsymbol{1}-\boldsymbol{D}\left(\boldsymbol{G}\left(\boldsymbol{\varepsilon}(\boldsymbol{X})\right)\right)\right)\right] \quad (1)$$

where $D(X)$ is the discriminator output, $E_{X\sim pX}$ represents the true sample sampled from the real space, $\log\left(D(X)\right)$ means that the original sample is expected to be judged as true by the discriminator, and $\log\left(1-D\left(G\left(\varepsilon(X)\right)\right)\right)$ means that the generated sample is expected to be considered false.

Feature Loss

GAN may lead to training instability when both the generator and the discriminator try to optimize the losses. To solve this problem, we introduced the feature matching proposed by Salimans et al. (2016) and used this loss function to stabilize the model training. It is defined by the following equation:

$$\boldsymbol{L_{fea}} = \boldsymbol{E_{X\sim pX}}\left\|\boldsymbol{f}(\boldsymbol{X}) - \boldsymbol{f}\left(\boldsymbol{G}\left(\boldsymbol{\varepsilon}(\boldsymbol{X})\right)\right)\right\|_{\boldsymbol{2}} \quad (2)$$

where $f(*)$ is the output of the last layer of the discriminator, and the loss is L_2 norm of (X) and $f\left(G\left(\varepsilon(X)\right)\right)$.

Mapping Loss

The goal of the model is to learn two mappings to reconstruct the sample. However, relying only on adversarial loss does not guarantee that a single original sample x_i can be mapped to the latent vector z_i and thus reconstructed as $\hat{x}_i$. To reduce the search space in the mapping process, we minimize the L_2 norm of residuals of the original and reconstructed samples. Its loss can be calculated as

$$\boldsymbol{L_{map}} = \boldsymbol{E_{X\sim pX}}\left\|\boldsymbol{X} - \boldsymbol{G}\left(\boldsymbol{\varepsilon}(\boldsymbol{X})\right)\right\|_{\boldsymbol{2}} \quad (3)$$

The generator tries to minimize the loss function, and the final overall loss function is obtained by combining (1), (2), and (3), as

$$\boldsymbol{L_G} = \boldsymbol{\lambda_a L_{adv}} + \boldsymbol{\lambda_f L_{fea}} + \boldsymbol{\lambda_m L_{map}} \quad (4)$$

where λ_a, λ_f, and λ_m represent the weights of each loss function, respectively.

The generator is trained directly using the adversarial loss in an attempt to maximize the following adversarial loss:

$$\boldsymbol{L_D} = \boldsymbol{E_{X\sim pX}}\left[\log\left(\boldsymbol{D}(\boldsymbol{X})\right)\right] + \boldsymbol{E_{X\sim pX}}\left[\log\left(\boldsymbol{1}-\boldsymbol{D}\left(\boldsymbol{G}\left(\boldsymbol{\varepsilon}(\boldsymbol{X})\right)\right)\right)\right] \quad (5)$$

The set loss function will be optimally searched by employing stochastic gradient descent (SGD). After continuous iterative adversarial learning, the discriminator and generator performance are gradually improved. When the set epoch or loss function convergence is reached, the GAN model can learn the distribution of normal data. After that, anomaly scores can be designed based on the model output to detect anomalies (described in Section "Anomaly Detection").

The Architecture of Generator and Discriminator

To improve the reconstruction effect, we optimize the design of the generator for time series data, and the basic model of the generator and discriminator of GAN is designated as LSTM. Inspired by the Jiang (2020), the LSTM module is improved into the Enhanced LSTM structure, which consists of multiple LSTM structures, as shown in **Figure 3**. It shows in detail the internal structure of the 3-layer Enhanced LSTM we introduced. The horizontal direction is the time step of the LSTM, that is, the time window size. The vertical is the number of LSTM layers, so that $h_t^{(0)}$ and $c_t^{(0)}$ are the hidden cell state and memory state of the first LSTM layer at moment t, respectively. The general LSTM passes the hidden cell state h_i and the memory state c_i horizontally to the LSTM cell at the next moment and passes h_i to the next LSTM layer in vertical direction.

Unlike the general stacked LSTM structure, the Enhanced LSTM binds both the hidden cell state and the memory state of each LSTM cell layer at a certain moment. It has the advantage of making full use of the hidden cell state and memory state of the current layer. The hidden cell state and the memory state of all layers except the current layer are used as the auxiliary input. This can improve the learning capability of the model network for time series data.

Recently, the attention mechanism has been widely applied in various research areas of neural networks. The attention mechanism allows the importance of different features to the final effect to be calculated, enabling the model to give higher weights to features that are beneficial to the outcome. In addition, the attention mechanism has a high degree of correctness and interpretability. RAIM (recurrent attentive and intensive model; Xu et al., 2018) was a model including an attention mechanism, which used multi-channel attention to improve the prediction of the model. Hashimoto et al. (2021) introduced RAIM into GAN to detect time series anomalies generated by semiconductor sensors. To this end, we consider a multi-channel attention mechanism, and an attention module is connected before both the encoder and the decoder.

The multi-channel attention module is divided into two stages, which can adaptively give different weights to multidimensional variables. The encoder and decoder structures

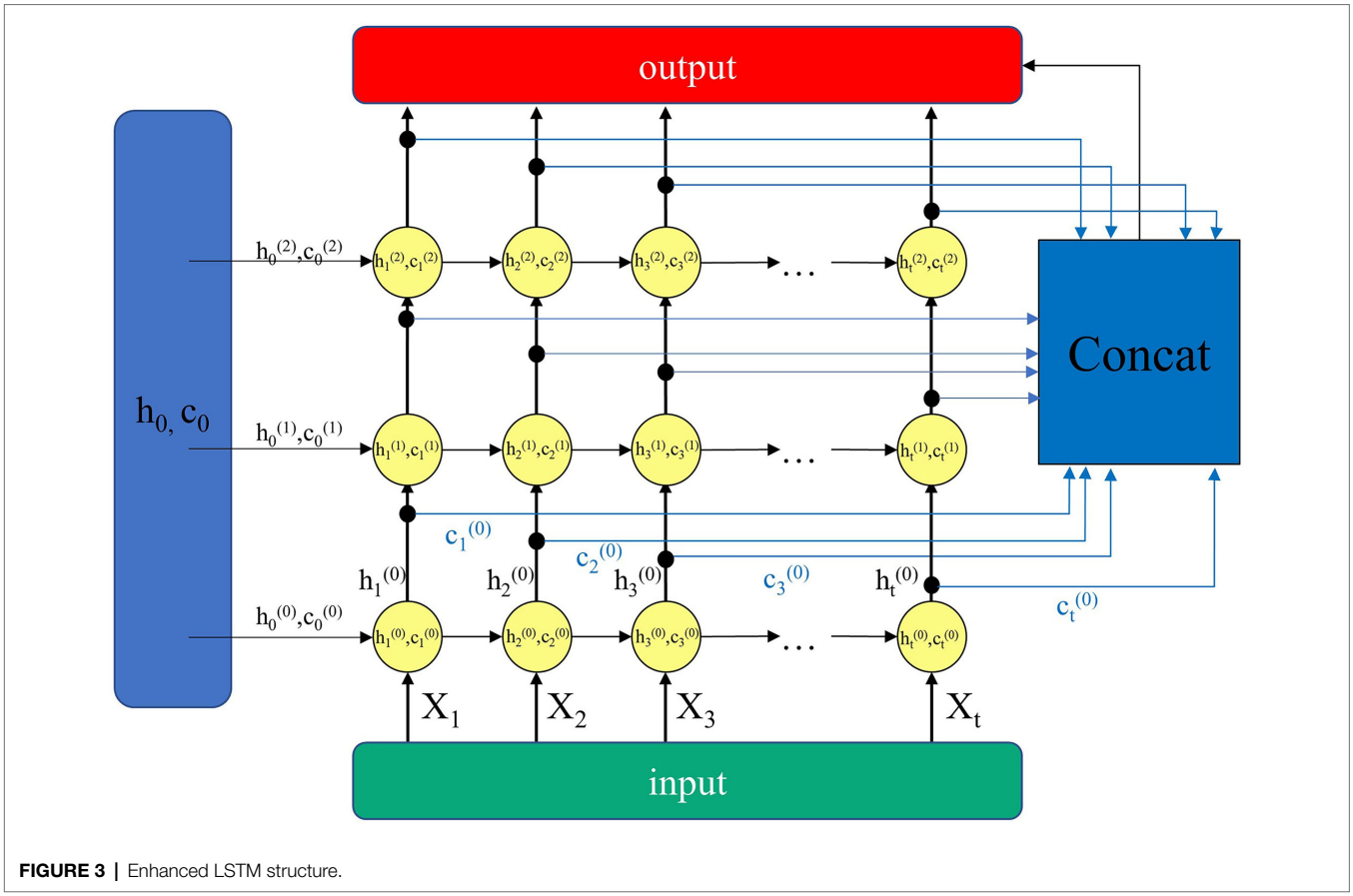


FIGURE 3 | Enhanced LSTM structure.

are shown in **Figures 4A,B**, respectively. The first stage performs the attention calculation in the time dimension. This mechanism based on the features extracted from the past sequences calculates the importance of different time steps of the input sequence. The input data $X = \{x_t, t = 1, \cdots, T\}$ is a series of time series data of length T, where the data $x_t \in \mathbb{R}^d$ at time t is an d-dimensional vector. Let the i-th subsequence be $X_i \in \mathbb{R}^{s_w \times d}$, when X is split by a sliding window s_w (Hashimoto et al., 2021). Then the importance a_{ij} of the time dimension is calculated by the following equation:

$$\boldsymbol{S_i^{time}} = \tanh\left(\boldsymbol{W_h^a} \cdot \boldsymbol{h_{i-1}} + \boldsymbol{X_i^T w_x^a} + \boldsymbol{b^a}\right) \tag{6}$$

$$\boldsymbol{a_{ij}} = \frac{\exp\left(\boldsymbol{S_{ij}^{time}}\right)}{\Sigma_{j'=1}^{\boldsymbol{S_w}} \exp\left(\boldsymbol{s_{ij'}^{time}}\right)}, \boldsymbol{j} = \boldsymbol{1, 2, \cdots, S_w} \tag{7}$$

where $W_h^a \epsilon \mathbb{R}^{s_w \times |h|}$ and $w_x^a \epsilon \mathbb{R}^{d \times 1}$ are the weighting matrices, $b^a \epsilon \mathbb{R}^{s_w \times 1}$ is the learning parameter through the attention mechanism, and h_{i-1} is the hidden state vector extracted from the Enhanced LSTM in the previous time step.

The second stage performs attention calculation in the feature dimension. Within the same time step, different weights are given according to the importance of different dimensional features, and the importance b_{ij} of each feature is calculated by the following equations:

$$\boldsymbol{S_i^{feature}} = \tanh\left(\boldsymbol{W_h^\beta} \cdot \boldsymbol{h_{i-1}} + \boldsymbol{X_i^T w_x^\beta} + \boldsymbol{b^\beta}\right) \tag{8}$$

$$\boldsymbol{b_{ij}} = \frac{\exp\left(\boldsymbol{S_{ik}^{feature}}\right)}{\Sigma_{k'=1}^{\boldsymbol{d}} \exp\left(\boldsymbol{s_{ik'}^{feature}}\right)}, \boldsymbol{k} = \boldsymbol{1, 2, \cdots, d} \tag{9}$$

where $W_h^\beta \epsilon \mathbb{R}^{d \times |h|}$ and $w_x^\beta \epsilon \mathbb{R}^{s_w \times 1}$ are the weighting matrices, $b^\beta \epsilon \mathbb{R}^{d \times 1}$ is the learning parameter through the attention mechanism, and h_{i-1} is the hidden state vector extracted from the Enhanced LSTM at the previous time step. After a two-stage attention mechanism, the model can capture important features more accurately.

The attention mechanism calculates the importance of the features and time of the input multidimensional time series data and weights them with the input to obtain a new input. The Enhanced LSTM captures the correlation and time dependence between the input data weighted by the attention mechanism. In the encoder, the Enhanced LSTM input is the original sample after weighting and the output is the

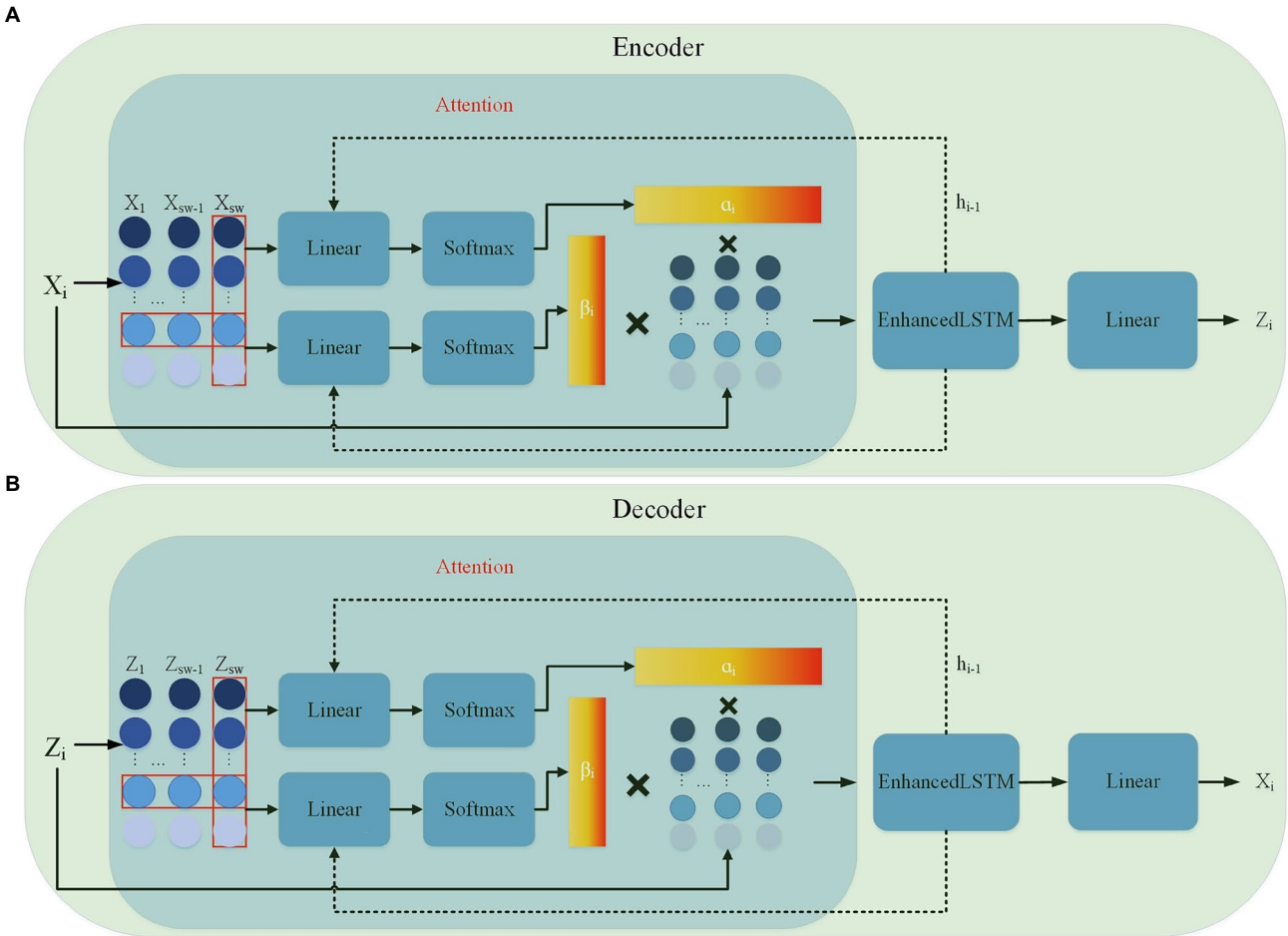


FIGURE 4 | Encoder–decoder internal detailed structure **(A)** encoder structure; **(B)** decoder structure.

feature vector. After that, the linear layer independently maps each feature vector as a latent vector each time. The decoder then reconstructs the latent vector into the original data by the same process. The discriminator is a simple Enhanced LSTM architecture, which is mainly used to distinguish the input samples as the real samples or the reconstructed samples. The performance of the discriminator and generator is gradually improved by adversarial learning.

Anomaly Detection

Our proposed model has been iteratively trained to learn the distribution pattern of normal data. The GAN model has the advantage of training a generator and a discriminator together, both of which can output metrics to help identify anomalies. The anomaly detection process is shown in **Figure 5**. The labeled test set data are divided into subsequences according to time windows using the same method as the training set data. The segmented time series data $X_{test} = \left\{ X_{test}^{j}, j = 1,2,\cdots n \right\} \subseteq R^{S_w \times N}$ will be binary classified. Each subsequence is determined to be normal (close to the normal data distribution) or abnormal (deviating from the normal data distribution) based on a threshold. We try different thresholds by using empirically determined threshold intervals and finally determine the threshold that results in optimal anomaly detection.

Reconstruction error is a measure of the difference between the true sample and the reconstructed sample. The ordinary generator only learns the mapping $G : Z \to X$ from random normal data in latent space to normal data, but there is no inverse mapping $G^{-1} : X \to Z$. That is to say, it is necessary to find the optimal latent vector $z \in Z$, such that the sample $G(z)$ reconstructed by the generator is closest to the test sample x_{test}^{j} in terms of distribution pattern. This process is the inversion of the generator. It needs to be further trained for the test sample to find the optimal latent vector, which generates the reconstruction sample with minimum error. The general procedure is to randomly sample $z_1 \in Z$ in the latent space and feed it into the generator to obtain the fake generative sequence $G(z_1)$. After that, the loss function is defined for the generated samples and the best latent vector z is found by gradient update in successive iterations. The degree of similarity

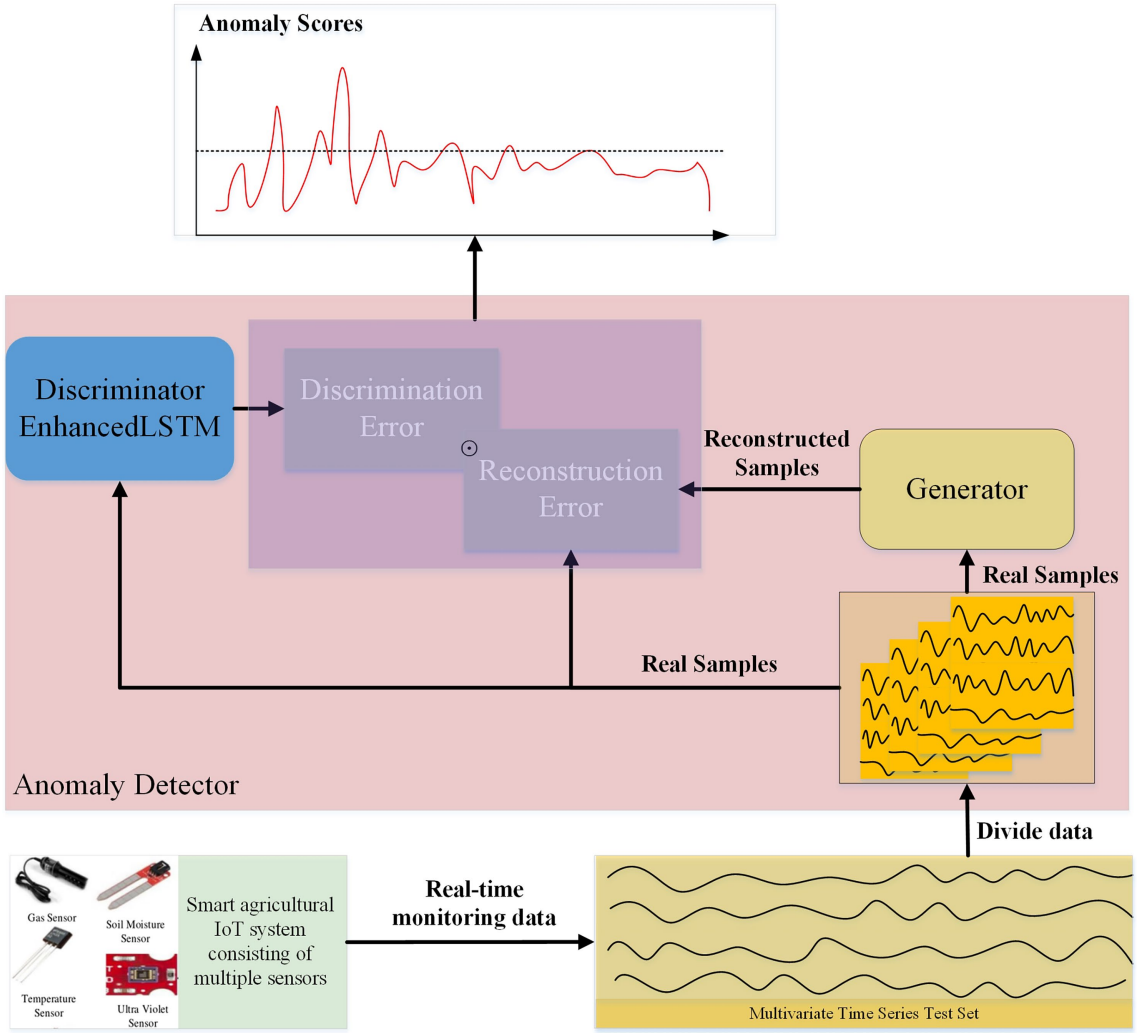


FIGURE 5 | Anomaly detection.

between the generated sample of the latent vector and the original sample determines the accuracy of the reconstruction error calculation.

In this paper, the generator has learned how to map the normal data in real space to the latent vector in latent space and then decode the latent vector back to the normal data. For the reconstructed samples, its corresponding latent vector is obtained by simply feeding it into the encoder without inversion. To improve the reconstruction effect and constrain the search domain, we add a new loss function (equation (3)). The reconstruction error can be obtained after the test data set is reconstructed by the generator. We combine two different error calculations to define the error more realistically.

The most intuitive way to measure the error is to use the point-wise difference, which directly calculates the difference between the corresponding points within each time step of the two series data. The error of the test data set at moment t is calculated as follows:

$$l_d = \sum_{i=1}^{n} \left| x_t^{test,i} - G\left(\varepsilon\left(x_t^{test,i}\right)\right) \right| \tag{10}$$

where $x_t^{test,i} \epsilon R^n$ is the measured value of i-th variables at moment t.

Time series data is a series of data points that make up a smooth curve. For this feature, we introduce the dynamic time warping (DTW) algorithm (Berndt and Clifford, 1994), which calculates the optimal match between two time series data and measures whether the two curves are similar in shape. This algorithm can solve the time shift issue of time series. As shown in **Figure 6**, there are two curves with the same shape, but they are not synchronized in time steps. In the actual error calculation, this should be determined as a low error. However, using a point-wise difference at the 10th time step leads to a larger error value. After the accumulation of multiple time steps, the error value may reach a level that affects the detection results. Based on this case, the DTW

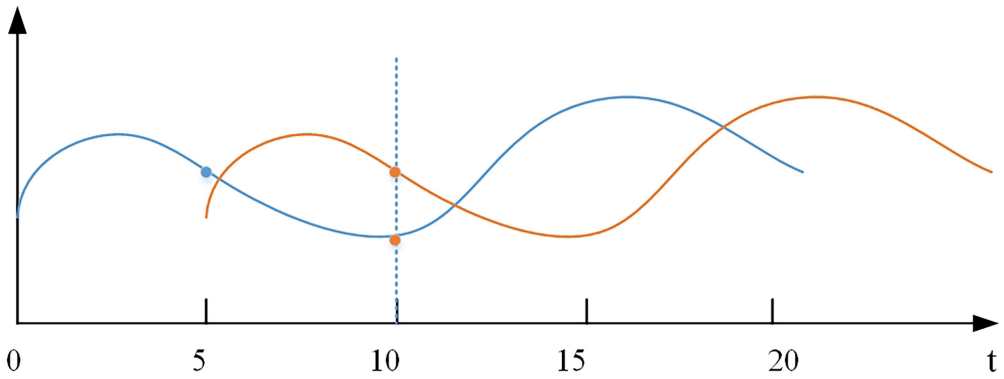


FIGURE 6 | Time shift issue of time series.

algorithm is introduced to measure the error more rationally. For the original subsequence $X=\left(x_t,x_{t+1},\ldots,x_{t+s_w-1}\right)$ and the reconstructed subsequence $\hat{X}=\left(\hat{x}_t,\hat{x}_{t+1},\ldots,\hat{x}_{t+s_w-1}\right)$, we define the matrix $W\epsilon R^{2*s_w\times 2*s_w}$, let the (i,j)-th element w_k represent the distance between x_i and $\hat{x}_j$. We want to find the warp path $w^*=\left(w_1{\cdot}w_2,\cdots,w_k\right)$ that defines the minimum distance between the two curves, subject to boundary conditions at the start and end, as well as constraints on continuity and monotonicity (Geiger et al., 2020). The two curve distances are defined as follows:

$$\boldsymbol{S_t}=\boldsymbol{W^*}=\boldsymbol{DTW}\left(\boldsymbol{X},\hat{\boldsymbol{X}}\right)=\min\left[\frac{\boldsymbol{1}}{\boldsymbol{k}}\sqrt{\sum_{\boldsymbol{k=1}}^{\boldsymbol{k}}\boldsymbol{w_k}}\right] \tag{11}$$

The final reconstruction error is given by using (10) and (11) as:

$$\boldsymbol{L_R}=\boldsymbol{\alpha L_d}+\boldsymbol{\beta S_t} \tag{12}$$

where α and β are the coefficients of the two reconstructed calculated values, which are the empirical values that make the experimental effect optimal.

During the training process, the main goal of the discriminator is to distinguish real samples from the generated samples and the output L_D (between 0 and 1) can be regarded as a parameter to determine whether the sequence is a real sample (close to 1) or a fake sample (close to 0). Thus, the output of the discriminator can be used as a measure of the anomaly score. The reconstruction error and the discriminator output are considered together as the final anomaly score. However, the reconstruction error and discriminator output cannot be simply used because a larger reconstruction error with a smaller discriminator output can lead to a very high anomaly score. The above problem is solved by using numerical normalization. The normalized result is calculated by the following formula to obtain the final anomaly score.

$$\boldsymbol{A}\left(\boldsymbol{x}\right)=\boldsymbol{\tau L_R}+\left(\boldsymbol{1}-\boldsymbol{\tau}\right)\boldsymbol{L_D} \tag{13}$$

where τ determines the relative importance of the two indicators (default value is 0.5).

Metrics, such as the precision of anomaly detection, can be calculated from the labeled test set data. Thresholds taken from the empirically determined threshold interval are used for anomaly detection. Different thresholds are obtained for different data sets, which results in optimal detection accuracy. Our proposed method is summarized in **Algorithm 1**.

EXPERIMENTS

In this section, we present the experimental design and the experimental results. The experimental design contains data set processing and parameter settings. The evaluation metrics for anomaly detection are briefly described. Experimental results include comparison experiments and model performance. Finally, the results of each experiment are discussed.

Datasets and Experimental Settings

To evaluate the performance of the proposed model, we test it on an agriculture-related time series data set. For future anomalies that may be encountered in smart agriculture, we mainly use three data sets that can represent relevant anomalous behaviors: SWMRU (USDA-ARS, 2016), KDDCUP99 (Blake and Merz, 1999), and HomeC (Taranveer, 2019).

The SWMRU data set contains 15-min mean weather data from the United States Department of Agriculture-Agricultural Research Service (USDA-ARS) Conservation and Production Laboratory (CPRL), Soil and Water Management Research Unit (SWMRU) research weather station, Bushland, Texas (Lat. 35.186714°, Long. -102.094189°, elevation 1,170 m above MSL) for all days in 2016. The data set has 18 variables and 35,139 time durations and it collects the values of sensors deployed at different heights on the grass during the irrigation season.

KDDCUP99 data set is the data set used for The Third International Knowledge Discovery and Data Mining Tools Competition. The data set is a network traffic data set that has 42 variables with 56,235 data points each. This data set is used to train a network intrusion detection model, which is adopted in our experiments to simulate possible anomalies in smart

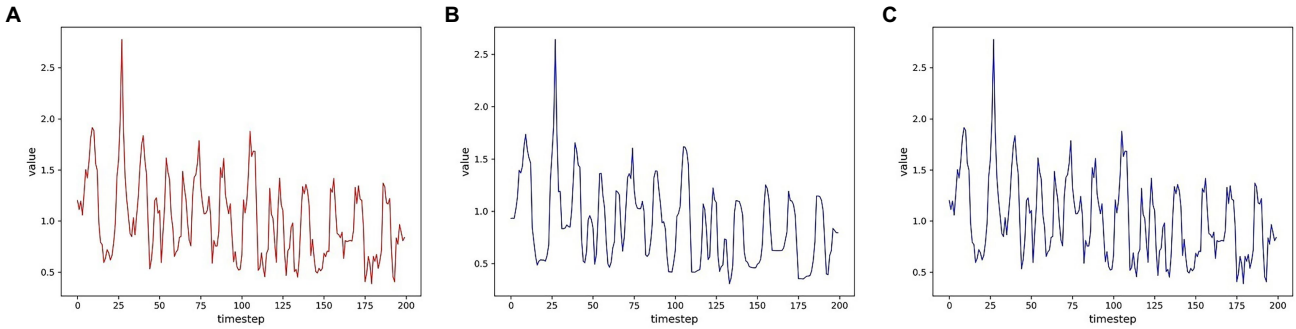


FIGURE 7 | Comparison of reconstruction effects on the SWMRU data set **(A)** Real sample; **(B)** No attention mechanism; **(C)** attention mechanism.

TABLE 1 | Details of datasets.

Dataset	Number of variables	The total length of time series	Proportion of anomaly
SWMRU	18	35,139	5%
KDDCUP99	42	56,235	19.5%
HomeC	32	503,900	8%

agricultural IoT due to intrusions. Most anomaly detection data sets have far fewer anomalous data points than normal data points, which leads to an imbalance in anomaly detection. This data set is a relatively balanced data set and its introduction allows for a more objective assessment of model performance.

The existing public data set for IoT power monitoring in agriculture has small dimensions, so the smart home IoT power usage data set HomeC can be used to simulate the power monitoring of IoT devices in future smart agriculture. It is collected in a smart home application scenario, and the data structure is similar to that of the agricultural IoT data set but with higher dimensionality. This data set contains 32 variables with 503,900 data points each. The information about the data set is presented in **Table 1**. The normal data points in the data set are marked as 1 and the abnormal data are marked as 0. The original data set is divided into the training set, the validation set, and the test set, and the training set contains only normal data. We use an unsupervised training approach, where the labeled validation set is used to find the optimal parameters of the model, and the labeled test set is used to compute the results of anomaly detection to evaluate the model performance.

For data preprocessing, we use a sliding window mechanism to partition the data as described in section 3.1. The optimal window size is an important element in the study of time series data. For this case, we use different window sizes, namely, $s_w = 30 \times i, i = 1, 2, \cdots 10$, to capture the state of the data at different accuracies. The results of this experiment are useful for exploring the effect of window size on detection performance. To better capture the normal data distribution, the training phase time step s_t is set to 10. During the testing phase, the time step is set to a time window size to ensure that anomalies are not repeatedly detected. The generator uses an Enhanced LSTM as the encoder and decoder, where the Enhanced LSTM depth is set to 3 and the hidden unit is set to 100. Generally, the discriminator follows the same parameter settings. However, unlike the generator, the final output dimension is 1, because the value of the discriminator indicates the degree of abnormality of the input sample. Li et al. (2019) evaluated the effect of latent vector dimensionality on the results in their experiments and verified that a dimension of 15 produced better data reconstruction, so we consider setting the dimension to 15 in our experiments. Since the discriminator converges faster, we set to train the discriminator once in one epoch but train the generator three times, with the epoch set to 100. The main parameters of the model are shown in **Table 2**.

Evaluation Measures

We use three standard evaluation measures, namely, Precision (Pre), Recall (Rec), and F1 score, to evaluate the anomaly detection performance of the proposed model:

$$Pre = \frac{TP}{TP + FP} \tag{14}$$

$$Rec = \frac{TP}{TP + FN} \tag{15}$$

$$F1 = 2 \times \frac{\mathrm{Pre} \times Rec}{Pre + Rec} \tag{16}$$

The objective of the model is anomaly detection, so the detected anomalies are positive samples. Therefore, TP is the correctly detected abnormal (True Positives: detected as abnormal while labeled as abnormal), FP denotes the incorrectly detected abnormal (False Positives: detected as abnormal while labeled as normal), TN represents the correctly detected normal (True Negatives: detected as normal while labeled as normal), and FN means the incorrectly detected normal (False Negatives: detected as normal while labeled as abnormal). $TP + FP$ denotes all the anomalies detected by the model, so precision indicates how many of the detected anomalies contain real anomalies,

TABLE 2 | Model parameter settings.

Window size	Training window step size	Test window step size	Input dimension	Number of LSTM hidden units	Number of LSTM layers	Latent space dimension
$30\times i, i=1,2,\cdots 10$	10	Window size	Data set dimension	100	3	15

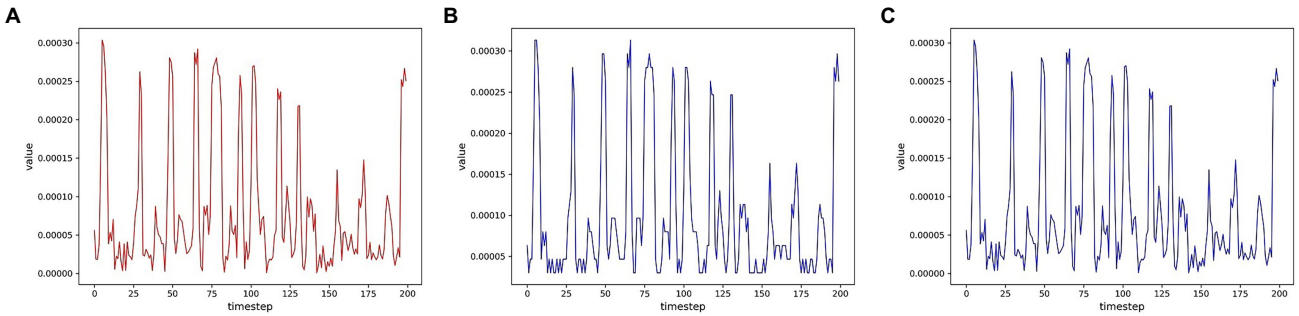


FIGURE 8 | Comparison of reconstruction effects on the HomeC data set **(A)** Real sample; **(B)** No attention mechanism; **(C)** Attention mechanism.

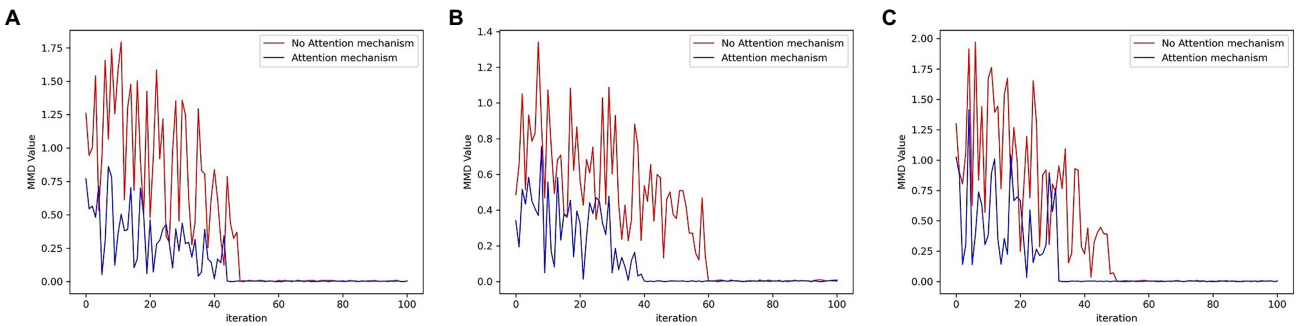


FIGURE 9 | MMD values for each data set **(A)** SWMRU; **(B)** KDDCUP99; **(C)** HomeC.

while $TP+FN$ is all the actual anomalies, so recall indicates how many of all the existing anomalies are detected by the model. The F1 score is the equal-weighted harmonic mean of the precision and recall. In the application scenario of anomaly detection anomalies are not common; that is, the distribution of anomalous and normal data is not balanced. Thus, the accuracy metric will not be used to evaluate the performance of the model.

Results and Discussion

We evaluate the anomaly detection performance of the proposed model on the above three data sets. To compare the performance of the models, MAD-GAN, TadGAN, TAnoGAN, and AutoEncoder (AE) were adopted to perform experiments on the same data set and record the experimental results. The above four counterpart models commonly used the reconstruction-based anomaly detection methods (Malhotra et al., 2016; Li et al., 2019; Bashar and Nayak, 2020; Geiger et al., 2020). In addition, to verify the validity of our proposed model structure, the results of the ablation experiments are shown and discussed.

Data Reconstruction Performance

To evaluate the reconstruction ability of the generator for the samples, we first visualize the multidimensional time series samples generated by the model with the original data. For more visualization, only one of these dimensions is shown for two data sets. In order to measure the degree of improvement of the attention mechanism on the reconstruction effect, the samples generated by the models without the attention mechanism are shown together. As shown in **Figures 7, 8**, the samples without the attention mechanism have largely conformed to the original sample distribution in terms of the overall trend. However, the comparison shows that the attention mechanism still leads to an improvement in the reconstruction effect. When the curve changes more dramatically, the generated samples are closer to the original samples because the attention mechanism allows the model to learn the samples more accurately.

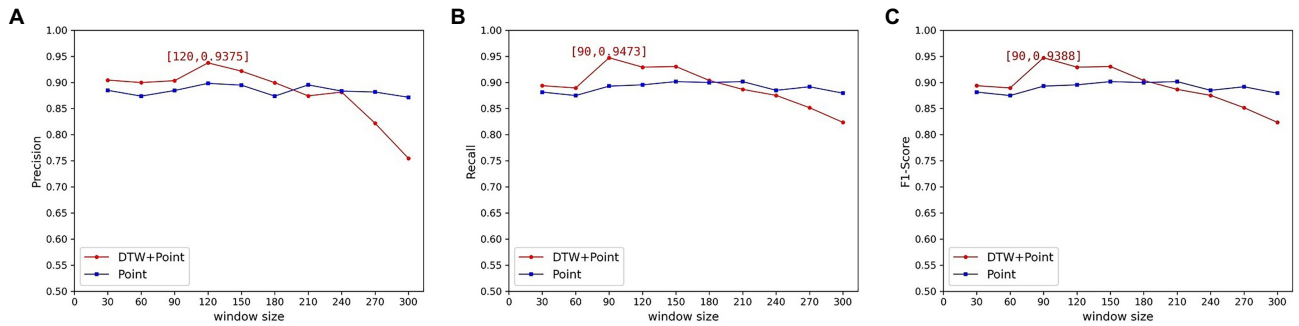


FIGURE 10 | Variation of metric with time window for the HomeC data set**(A)** Precision; **(B)** Recall; **(C)** F1 Score.

Input: real data X, test data Y.

Output: Anomaly scores of test data A.

Function *train* (X, m, *epochs*, *G_round*)**:**

m ←batch size

epochs ←number of iterations over the data

G_round ←number of iterations of the generator per epoch

for *epochs* **do**

Sample m vectors $\{x_1, \cdots, x_m\}$ from the real data

Reconstruct real data: $\boldsymbol{X} = \{\boldsymbol{x_i}\}_{i=1}^{m} \rightarrow \boldsymbol{G}(\boldsymbol{\varepsilon}(\boldsymbol{X}))$

Input X to the discriminator: $\boldsymbol{X} = \{\boldsymbol{x_i}\}_{i=1}^{m} \rightarrow \boldsymbol{D}(\boldsymbol{X})$

Input $G(\varepsilon(X))$ to the discriminator: $\boldsymbol{G}(\boldsymbol{\varepsilon}(\boldsymbol{X})) \rightarrow \boldsymbol{D}\left(\boldsymbol{G}(\boldsymbol{\varepsilon}(\boldsymbol{X}))\right)$

Update the discriminator parameters by maximizing L_D:

$$\boldsymbol{max}\frac{\boldsymbol{1}}{\boldsymbol{m}}\sum_{\boldsymbol{i=1}}^{\boldsymbol{m}}[\boldsymbol{log}(\boldsymbol{D}(\boldsymbol{x_i})) + \boldsymbol{log}(\boldsymbol{1} - \boldsymbol{D}(\boldsymbol{G}(\boldsymbol{\varepsilon}(\boldsymbol{x_i}))))]$$

for *G_round* **do**

Reconstruct real data: $\boldsymbol{X} = \{\boldsymbol{x_i}\}_{i=1}^{m} \rightarrow \boldsymbol{G}(\boldsymbol{\varepsilon}(\boldsymbol{X}))$

Update the generator parameters by minimizing L_G:

$$\boldsymbol{min}\frac{\boldsymbol{1}}{\boldsymbol{m}}\sum_{\boldsymbol{i=1}}^{\boldsymbol{m}}[\boldsymbol{\lambda_a L_{adv}} + \boldsymbol{\lambda_f L_{fea}} + \boldsymbol{\lambda_m L_{map}}]$$

end for

end for

return G, D

Function *anomalyScore* (Y, G, D)**:**

Sample n vectors $\{y_1, \cdots, y_n\}$ from the test data

Reconstruct test data: $\boldsymbol{Y} = \{\boldsymbol{y_i}\}_{i=1}^{n} \rightarrow \boldsymbol{G}(\boldsymbol{\varepsilon}(\boldsymbol{Y}))$=$\boldsymbol{\hat{Y}}$

Calculate the reconstruction error: $\boldsymbol{L_R} = \boldsymbol{f}(\boldsymbol{Y}, \boldsymbol{\hat{Y}})$

Calculating discriminator results: $\boldsymbol{G}(\boldsymbol{\varepsilon}(\boldsymbol{Y})) \rightarrow \boldsymbol{D}\left(\boldsymbol{G}(\boldsymbol{\varepsilon}(\boldsymbol{Y}))\right) = \boldsymbol{L_D}$

Calculate the Anomaly Score $\boldsymbol{A}(\boldsymbol{y}) = \boldsymbol{\tau L_R} + (\boldsymbol{1} - \boldsymbol{\tau})\boldsymbol{L_D}$

return A

ALGORITHM 1 | Algorithm for proposed method.

In addition, Maximum Mean Discrepancy (MMD) was used to evaluate whether the GAN model actually learns the distribution of the training data (Li et al., 2015, 2019). Therefore, MMD was also introduced into the experiment to compare the effect of reconstruction. The decrease in the MMD values indicates that the data generated by the model conform more to the distribution of the original sample. The MMD values generated from the three data sets by iterative training of the GAN are plotted in **Figure 9**. As shown in these figures, as the number of iterations increases, the model outputs samples that are increasingly closer to the original samples. And the samples generated in the three data sets by the model incorporating the attention mechanism obtained lower MMD values. The MMD value more clearly illustrates that the attention mechanism improves the reconfiguration effect.

Window Setting and Reconstruction Error Metric

The sliding window size setting is critical to the processing of time series data, so we conduct experiments on the validation data set to determine the appropriate window size. The relationship between the sliding window size setting and the reconstruction error metric will also be discussed in this section. Next, the experiments are described using the HmoeC data set as an example. The sliding window size is still set to $s_W = 30 \times i, i = 1, 2, \cdots 10$, but the reconstruction error is calculated in two ways to explore its relationship with the window size. The model proposed in this paper uses the point-wise difference coupled with the DTW algorithm results as the final reconstruction error calculation, where the parameters α and β are derived from multiple experiments on the validation set tuned according to different data sets.

In the previous experiments, we used the coupling results as a reconstruction error metric to determine its potential correlation with the window size. It was found experimentally that all three indicators of the experiment showed a decreasing trend as the time window increased. And when the time window increases to a certain extent, these indicators show a large decline. However, the common models that use point-wise error as a reconstruction metric do not show this phenomenon. For comparison, we conducted experiments using the universal point-wise difference calculation (β = 0). None of the three indicators showed a significant decrease with increasing time windows. The precision, recall, and F1 scores of the data set are shown in **Figure 10**.

From the above experimental results, it can be seen that the best values of precision, recall, and F1 score can indeed be obtained by using the calculation method proposed in this paper. However, as the time window increased, the experimental results showed a significant downward trend. This is because the DTW algorithm outputs the similarity measure of the two curves, and it is proposed to solve the time shift problem of the curves. But the tolerance of the DTW algorithm to curve similarity increases due to the excessive time window. The DTW algorithm has a small probability of finding the optimal distance between two curves in a small window and outputs a large calculated value. The anomaly will be detected because it improves the reconstruction error value. When the window becomes larger, the DTW algorithm can always find the corresponding minimum distance. Therefore, the output value is reduced, resulting in the abnormal subsequence being incorrectly classified as normal.

The experimental results also clarify that the window size does have a large effect on the results, so the window size needs to be determined reasonably. The window size should be strictly determined when using DTW as a reconstruction error metric. The experimental time window size is finally determined to be 90, at which time the optimal F1 score is 0.9388.

Comparison Experiments and Discussion

The window size is set to 90 based on the results obtained in section 4.3.2, and the remaining hyperparameters are tuned using the validation set. Anomaly detection is performed on three test data sets using our proposed model with optimal hyperparameters to obtain precision, recall values, and F1 scores. To demonstrate the effectiveness of our proposed model, we conduct experiments using each of the four reconstruction-based anomaly detection models mentioned above, including MAD-GAN, Tad-GAN, TAnoGAN, and AE. The average values of Precision, Recall, and F1 score after ten rounds are calculated on three data sets, and the comparison results of five algorithms are shown in **Table 3**. As shown in **Table 3**, the metrics of our proposed model exceed 0.9 on all three data sets and outperform other algorithms in several of the three metrics. The experimental results indicate that the model proposed in this paper has better performance and outperforms other algorithms in the specified data set.

In addition, to demonstrate the validity of our proposed model improvement, the results of the ablation experiments are also presented in **Table 3**. To prove the effectiveness of the encoder–decoder architecture containing the attention mechanism, we modify the generator to the same LSTM architecture as MAD-GAN. At the same time, the discriminator and the error calculation method are kept constant. The reconstruction error is obtained after generator inversion. The experimental results are noted as "Ours-Gen," that is, the experimental results obtained by removing the improvements of the generator. To demonstrate the boosting effect of Enhanced LSTM on the discriminator, we also keep the remaining architecture constant and only change the generator to the general LSTM architecture. Since there is no change in the generator, the experimental results can be obtained directly without generator inversion. The experimental results are noted as "Ours-Dis," which is the experimental result obtained after removing the discriminator improvement. For the error calculation method, the experimental results are noted as "Ours-Gen-Dis." After the generator and discriminator improvements are all removed, the remaining architecture of our proposed model is equivalent to MAD-GAN except for the error calculation method. As shown in **Table 3**, after removing the generator improvements, our proposed model shows a substantial decrease in the experimental metrics for all three data sets. It is concluded that our proposed encoder–decoder architecture incorporating the attention mechanism does improve the model performance. In addition, the results of the "Ours-Dis" also showed a small decrease. The Enhanced LSTM that was introduced into the discriminator is also relevant for model performance improvement. The experimental metrics of "Ours-Gen-Dis" are higher than MAD-GAN, which can prove that our proposed error calculation method improves the detection effect.

The model proposed in this paper significantly outperforms AE, MAD-GAN, and TAnoGAN in all three metrics. Our proposed model generator is similar to AE, but the final detection performance is better than AE. The autoencoder alone does not detect anomalies very well, because the autoencoder trained with appropriate loss functions in adversarial training is better able to learn the general data distribution. The better the generator learns normal data, the more sensitive it is for the abnormal data in anomaly detection. For MAD-GAN and TAnoGAN, these two models share a similar structure, in which their generators are similar to simple decoders. They both use random normal data directly to generate the reconstruction data, after which the reconstruction error is calculated. In order to obtain the accurate reconstruction error, it is necessary to find its

TABLE 3 | Experimental results of different methods on three data sets.

Data set	Methods	Precision (%)	Recall (%)	F1 score
SWMRU	Ours	**92.37**	**95.55**	**0.9482**
	Ours-Dis	92.05	94.50	0.9403
	Tad-GAN	91.08	94.13	0.9348
	Ours-Gen	87.95	90.31	0.8891
	Ours-Gen-Dis	87.16	89.94	0.8835
	MAD-GAN	85.41	89.23	0.8754
	TAnoGAN	86.43	89.35	0.8876
	AE	69.48	75.26	0.7238
KDDCUP99	Ours	**93.51**	**96.25**	0.9385
	Ours-Dis	93.38	96.05	0.9365
	Tad-GAN	93.17	94.83	**0.9405**
	Ours-Gen	87.19	92.35	0.8847
	Ours-Gen-Dis	86.49	91.42	0.8794
	MAD-GAN	83.65	89.30	0.8689
	TAnoGAN	85.58	88.23	0.8736
	AE	75.43	81.41	0.7749
HomeC	Ours	**91.12**	**92.79**	**0.9235**
	Ours-Dis	90.73	92.48	0.9205
	Tad-GAN	88.49	92.53	0.9199
	Ours-Gen	85.74	88.63	0.8749
	Ours-Gen-Dis	84.16	88.11	0.8636
	MAD-GAN	83.39	87.09	0.8419
	TAnoGAN	82.24	88.67	0.8676
	AE	67.26	72.84	0.6929

The bold values in Table 3 are the highest values of each experimental metric for each data set.

corresponding optimal latent vector for the test sample. In the reconstruction error calculation process of the two models mentioned above, the best latent vector is derived from the inversion of the test sample by the generator. This may allow the model to improve the reconstruction performance on the test sample, thus allowing the reconstruction error values to be reduced to the extent that affects the final test results. Our models are trained based on normal samples, and both the encoder and decoder learn the two mappings based on the distribution pattern of the normal samples. The encoder learns how to map a normal sample to a latent vector to reconstruct the normal sample. And for test samples that may have anomalies, the encoder mapping may lose some information. The same is true for the decoder, which learning goal is to improve the reconstruction ability of the latent vector for normal samples. After two mappings, test samples with distribution patterns that differ significantly from the normal sample may yield greater reconstruction errors. In other words, the encoder–decoder structure can widen the gap between normal and abnormal samples, which helps to improve detection performance.

On the other hand, TadGAN introduced the cycle-consistent loss and trained the encoder together with the generator, which was used to learn the mapping of normal data to latent vectors. Both this model and our model train the encoder and generator together, so they have almost similar experimental performance. The difference is that this model used cycle-consistent loss for training and introduced a new discriminator for the encoder to improve learning, whereas our model improves learning through an attention mechanism. Both training methods prevent the contradiction between the encoder and the generator and find the corresponding optimal latent vector to the test sample using the most direct method. TadGAN explored different ways of coupling different reconfiguration computations with discriminator outputs, and we have conducted experiments using its best structure. The average F1 score of this model is higher than that of our model, but the recall of our model is higher than it. This proves that our model can detect more anomalies that are present actually. Meanwhile, the optimal F1 value of our model outperformed it in ten training rounds.

CONCLUSION

In this paper, we proposed a GAN-based anomaly detection model for multidimensional time series data generated in smart agricultural IoT. This model used the GAN architecture to learn the distribution patterns of normal data and applied reconstruction methods for anomaly detection. Considering the time dependence of time series data and the potential correlation between multidimensional variables, an improved Enhanced LSTM network to form the basis of the GAN was considered in this model. For the problem of generator inversion, the encoder–decoder architecture was adopted as the generator structure of GAN. The co-training of the encoder and decoder eliminated the inversion of the generator for test samples. This effectively reduced the computation time and met the demand for real-time anomaly detection. The performance of anomaly detection has been improved by the use of encoder–decoder architecture. To further improve the reconstruction effect, the encoder–decoder architecture incorporates an attention mechanism, which can extract weights in the time and feature dimensions to help the model reconstruct the samples. For anomaly detection, a new anomaly score calculation was proposed, which took the coupled result of the point-wise difference error and the curve similarity metric as the reconstruction error. The point-wise error and curve similarity were considered together to better fit the definition of realistic error.

Experiments were designed on three smart agriculture-related data sets and these results were compared with four previous anomaly detection algorithms to verify the effectiveness and superiority of the algorithm. The results proved that our method outperformed other methods in most of the metrics, and the error calculation method proposed in this paper can better detect the anomaly. Not only that, our proposed model obtained superior experimental metrics on high-dimensional smart agriculture data sets, which also reflects that GAN can better handle high-dimensional time series data. With the continuous development of smart agriculture, the dimensionality and quantity of data will grow. The model proposed in this paper also provides a new and useful insight for the anomaly detection of high-dimensional time series data in smart agriculture. However, the time window size setting needed to be considered primarily, which may be the reason why this model is lower than one of the counterpart models in terms of F1 score. Thus, how to choose time windows in time series is an important research topic, and the calculation method proposed in this paper also has a strong correlation with the size of time windows, we will continue our work on anomaly calculation methods and time windows in the future.

DATA AVAILABILITY STATEMENT

Publicly available datasets were analyzed in this study. These data can be found at: https://archive.ics.uci.edu/ml/datasets/KDD+Cup+1999+Data, https://www.kaggle.com/taranvee/smart-home-dataset-with-weather-information, and https://catalog.data.gov/dataset/data-from-quality-controlled-research-weather-data-usda-ars-bushland-texas.

AUTHOR CONTRIBUTIONS

WC: conceptualization, methodology, and supervision. TM: model and experiment design and analysis and writing—original draft preparation. XW: software and writing—editing. GW: simulation, writing—review, and project administration. All authors have read and agreed to the published version of the manuscript.

FUNDING

This research was funded in part by the Open Project Fund of Key Laboratory of Mine Disaster Prevention and Control under grant no. SMDPC202102 and supported in part by the Graduate Research and Practice Projects of Minzu University of China.

REFERENCES

Abdallah, M., Lee, W. J., Raghunathan, N., Sutherland, J. W., and Bagchi, S. (2021). Anomaly detection through transfer learning in agriculture and manufacturing IoT systems. *arXiv.* [Preprint]. doi: 10.48550/arXiv.2102.05814

Adhikari, R., and Agrawal, R. K. (2013). An introductory study on time series modeling and forecasting. *arXiv.* [Preprint]. doi: 10.48550/arXiv.1302.6613

Adkisson, M., Kimmell, J. C., Gupta, M., et al. (2021). "Autoencoder-based anomaly detection in smart farming ecosystem," In *2021 IEEE International Conference on Big Data*; December 15, 2021.

Aggarwal, C. C. (2017). *An Introduction to Outlier Analysis*. Cham: Springer.

Akçay, S., Atapour-Abarghouei, A., and Breckon, T. P. (2019). "Skip-ganomaly: skip connected and adversarially trained encoder-decoder anomaly detection." in *2019 International Joint Conference on Neural Networks*; July 14, 2019.

Bashar, M. A., and Nayak, R. (2020). "TAnoGAN: time series anomaly detection with generative adversarial networks." in *2020 IEEE Symposium Series on Computational Intelligence*; December 1, 2020.

Berndt, D. J., and Clifford, J. (1994). Using dynamic time warping to find patterns in time series. *KDD Workshop* 10, 359–370.

Blake, C., and Merz, C. J. (1999). Data from: The third international knowledge discovery and data mining tools competition. UCI repository of machine learning databases. Available at: https://archive.ics.uci.edu/ml/datasets/KDD+Cup+1999+Data (Accessed December 10, 2021).

Cao, X., Yao, Y., Li, L., Zhang, W., An, Z., Zhang, Z., et al. (2021). iGrow: A smart agriculture solution to autonomous greenhouse control. *arXiv.* [Preprint]. doi: 10.48550/arXiv.2107.05464

Chalapathy, R., and Chawla, S. (2019). Deep learning for anomaly detection: a survey. *arXiv.* [Preprint].

Cook, A. A., Mısırlı, G., and Fan, Z. (2019). Anomaly detection for IoT time-series data: A survey. *IEEE Internet Things J.* 7, 6481–6494. doi: 10.1109/JIOT.2019.2958185

Farooq, M. S., Riaz, S., Abid, A., Abid, K., and Naeem, M. A. (2019). A survey on the role of IoT in agriculture for the implementation of smart farming. *IEEE Access.* 7, 156237–156271. doi: 10.1109/ACCESS.2019.2949703

Garg, S., Pundir, P., Jindal, H., Saini, H., and Garg, S. (2021). "Towards a multimodal system for precision agriculture using IoT and machine learning." in *2021 12th international conference on computing communication and networking technologies (ICCCNT)*; July 6, 2021.

Geiger, A., Liu, D., Alnegheimish, S., et al. (2020). "TadGAN: time series anomaly detection using generative adversarial networks." in *2020 IEEE International Conference on Big Data*; December 10, 2020.

Goodfellow, I., Pouget-Abadie, J., Mirza, M., Xu, B., Warde-Farley, D., Ozair, S., et al. (2014). *Generative Adversarial Nets. Advances in Neural Information Processing Systems*. Red Hook: NY Curran.

Hasan, M., Islam, M. M., Zarif, M. I. I., and Hashem, M. M. A. (2019). Attack and anomaly detection in IoT sensors in IoT sites using machine learning approaches. *Internet Things.* 7:100059. doi: 10.1016/j.iot.2019.100059

Hashimoto, M., Ide, Y., and Aritsugi, M. (2021). Anomaly detection for sensor data of semiconductor manufacturing equipment using a GAN. *Procedia Comp. Sci.* 192, 873–882. doi: 10.1016/j.procs.2021.08.090

Hawkins, D. M. (1980). *Identification of Outliers. Vol. 1*. London: Chapman and Hall.

Jiang, Z. K. (2020). A Deep Learning Framework for IoT Time-Series Data Analysis and Applications. Master's Thesis. Changsha (Hunan): Hunan University.

Khalil, R. A., Saeed, N., Masood, M., Fard, Y. M., Alouini, M. S., and Al-Naffouri, T. Y. (2021). Deep learning in the industrial internet of things: potentials, challenges, and emerging applications. *IEEE Internet Things J.* 8, 11016–11040. doi: 10.1109/JIOT.2021.3051414

Kwon, D., Kim, H., Kim, J., Suh, S. C., Kim, I., and Kim, K. J. (2019). A survey of deep learning-based network anomaly detection. *Clust. Comput.* 22, 949–961. doi: 10.1007/s10586-017-1117-8

Li, D., Chen, D., Goh, J., and Ng, S. K. (2018). Anomaly detection with generative adversarial networks for multivariate time series. *arXiv*. [Preprint]. doi: 10.48550/arXiv.1809.04758

Li, D., Chen, D., Jin, B., Shi, L., and Goh, J., (2019). "MAD-GAN: multivariate anomaly detection for time series data with generative adversarial networks," in *International Conference on Artificial Neural Networks*; September 17, 2019.

Li, Y., Swersky, K., and Zemel, R. (2015). "Generative moment matching networks." in *International Conference on Machine Learning*. 1718–1727. PMLR.

Li, W., Wang, D., and Li, M., Gao Gao, Y., Wu, J., and Yang, X., (2021). Field detection of tiny pests from sticky trap images using deep learning in agricultural greenhouse. *Comput. Electron. Agric.* 183:106048. doi:10.1016/j.compag.2021.106048

Liu, J., and Wang, X. (2020). Tomato diseases and pests detection based on improved Yolo V3 convolutional neural network. *Front. Plant Sci.* 11:898. doi: 10.3389/fpls.2020.00898

Malhotra, P., Ramakrishnan, A., Anand, G., et al. (2016). LSTM-based encoder-decoder for multi-sensor anomaly detection. *arXiv*. [Preprint]. doi: 10.48550/arXiv.1607.00148

Moso, J. C., Cormier, S., De Runz, C., Fouchal, H., Wandeto, J. M., et al. (2021). Anomaly detection on data streams for smart agriculture. *Agriculture* 11:1083. doi: 10.3390/agriculture11111083

Nassif, A. B., Talib, M. A., Nasir, Q., and Dakalbab, F. M. (2021). Machine learning for anomaly detection: a systematic review. *IEEE Access.* 9, 78658–78700. doi: 10.1109/ACCESS.2021.3083060

Pang, G., Shen, C., Cao, L., and Hengel, A. V. D. (2021). Deep learning for anomaly detection: A review. *ACM Comp. Surveys* 54, 1–38. doi: 10.1145/3439950

Park, H., Park, D., and Kim, S. (2021). Anomaly detection of operating equipment in livestock farms using deep learning techniques. *Electronics* 10:1958. doi: 10.3390/electronics10161958

Pundir, M., and Sandhu, J. K. (2021). A systematic review of quality of service in wireless sensor networks using machine learning: recent trend and future vision. *J. Netw. Comput. Appl.* 188:103084. doi: 10.1016/j.jnca.2021.103084

Rafii, F., and Kechadi, T. (2019). "Collection of historical weather data: issues with missing values," in *Proceedings of the 4th International Conference on Smart City Applications*; October 2, 2019.

Salimans, T., Goodfellow, I., Zaremba, W., Cheung, V., Radford, A., and Chen, X. (2016). Improved techniques for training gans. *arXiv.* [Preprint]. doi: 10.48550/arXiv.1606.03498

Schlegl, T., Seeböck, P., Waldstein, S. M., Schmidt-Erfurth, U., and Langs, G. (2017). "Unsupervised anomaly detection with generative adversarial networks to guide marker discovery." in *International Conference on Information Processing in Medical Imaging*; June 25, 2017.

Taranveer (2019) Data from: The readings with a time span of 1 minute of house appliances in kW from a smart meter and weather conditions of that particular region. Kaggle. Available at: https://www.kaggle.com/taranvee/smart-home-dataset-with-weather-information (Accessed January 10, 2022).

USDA-ARS (2016). Data from: The USDA-ARS conservation and production laboratory (CPRL), soil and water management research unit (SWMRU) research weather station. Data.Gov. Available at: https://catalog.data.gov/dataset/data-from-quality-controlled-research-weather-data-usda-ars-bushland-texas (Accessed January 12, 2022).

Vilenski, E., Bak, P., and Rosenblatt, J. D. (2019). Multivariate anomaly detection for ensuring data quality of dendrometer sensor networks. *Comput. Electron. Agric.* 162, 412–421. doi: 10.1016/j.compag.2019.04.018

Vyas, A., and Bandyopadhyay, S. (2020). Semi-supervised soil moisture prediction through graph neural networks. *arXiv*. [Preprint]. doi: 10.48550/arXiv.2012.03506

Xu, Y., Biswal, S., Deshpande, S. R., et al. (2018). "Raim: recurrent attentive and intensive model of multimodal patient monitoring data." in *Proceedings of the 24th ACM SIGKDD international conference on Knowledge Discovery and Data Mining*; July 19, 2018.

Yang, Y., Ding, S., Liu, Y., Meng, S., Chi, X., Ma, R., et al. (2021). Fast wireless sensor for anomaly detection based on data stream in an edge-computing-enabled smart greenhouse. *Digital Comm. Networks* doi: 10.1016/j.dcan.2021.11.004 [Epub ahead of Print].

Yu, J., Song, Y., Tang, D., Han, D., and Dai, J. (2021). Telemetry data-based spacecraft anomaly detection with spatial–temporal generative adversarial networks. *IEEE Trans. Instrum. Meas.* 70, 1–9. doi: 10.1109/TIM.2021.3073442

Zenati, H., Foo, C. S., Lecouat, B., Manek, G., and Chandrasekhar, V. R. (2018a). Efficient Gan-based anomaly detection. *arXiv*. [Preprint]. doi: 10.48550/arXiv.1802.06222

Zenati, H., Romain, M., Foo, C. S., Lecouat, B., and Chandrasekhar, V. (2018b). "Adversarially learned anomaly detection," in *2018 IEEE International Conference on Data Mining*; November 17, 2018.

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

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

frontiers | Frontiers in Plant Science

ORIGINAL RESEARCH
published: 09 June 2022
doi: 10.3389/fpls.2022.884891

Circle Fitting Based Image Segmentation and Multi-Scale Block Local Binary Pattern Based Distinction of Ring Rot and Anthracnose on Apple Fruits

*Qin Feng[1], Shutong Wang[2], He Wang[3], Zhilin Qin[1] and Haiguang Wang[1]**

[1] College of Plant Protection, China Agricultural University, Beijing, China, [2] College of Plant Protection, Hebei Agricultural University, Baoding, China, [3] Forest Pest Management and Quarantine Station of Beijing, Beijing, China

OPEN ACCESS

Edited by:
Lei Shu,
Nanjing Agricultural University, China

Reviewed by:
Chuanlei Zhang,
Tianjin University of Science and Technology, China
Chenliang Yu,
Zhejiang Agriculture and Forestry University, China

***Correspondence:**
Haiguang Wang
wanghaiguang@cau.edu.cn

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

Received: *27 February 2022*
Accepted: *11 May 2022*
Published: *09 June 2022*

Citation:
Feng Q, Wang S, Wang H, Qin Z and Wang H (2022) Circle Fitting Based Image Segmentation and Multi-Scale Block Local Binary Pattern Based Distinction of Ring Rot and Anthracnose on Apple Fruits. Front. Plant Sci. 13:884891. doi: 10.3389/fpls.2022.884891

Ring rot caused by *Botryosphaeria dothidea* and anthracnose caused by *Colletotrichum gloeosporioides* are two important apple fruit diseases. It is critical to conduct timely and accurate distinction and diagnosis of the two diseases for apple disease management and apple quality control. The automatic distinction between the two diseases was investigated based on image processing technology in this study. The acquired disease images were preprocessed via image scaling, color image contrast stretching, and morphological opening and closing reconstruction. Then, two lesion segmentation methods based on circle fitting were proposed and used to conduct lesion segmentation. After comparison with the manual segmentation results obtained via the software Adobe Photoshop CC, Lesion segmentation method 1 was chosen for further disease image processing. The gray images on the nine components in the RGB, HSI, and L*a*b* color spaces of the segmented lesion images were filtered by using multi-scale block local binary pattern operators with the sizes of pixel blocks of 1×1, 2×2, and 3×3, respectively, and the corresponding local binary pattern (LBP) histogram vectors were calculated as the features of the lesion images. Subsequently, support vector machine (SVM) models and random forest models were built based on individual LBP histogram features or different LBP histogram feature combinations for distinguishing the diseases. The optimal SVM model with the distinction accuracies of the training and testing sets equal to 100 and 95.12% and the optimal random forest model with the distinction accuracies of the training and testing sets equal to 100 and 90.24% were achieved. The results indicated that the distinction between the two diseases could be implemented with high accuracy by using the proposed method. In this study, a method based on image processing technology was provided for the distinction of ring rot and anthracnose on apple fruits.

Keywords: apple ring rot, apple anthracnose, image distinction, circle fitting, multi-scale block local binary pattern, support vector machine, random forest

INTRODUCTION

Apple is a kind of fruit with great commercial value, and it is an important kind of export fruit in China (Chen et al., 2010). Apple ring rot caused by *Botryosphaeria dothidea* and apple anthracnose caused by *Colletotrichum gloeosporioides* is two common diseases on apple fruits (Li B. H. et al., 2013; Hu et al., 2016). These two kinds of diseases form lesions on the apple fruit surface and cause decay on apple fruits, resulting in severe yield losses and quality declines of apple fruits. Lesions on apple fruits caused by ring rot are usually very similar to those caused by anthracnose. Agricultural technicians with rich practical experience are required to differentiate and identify the two apple fruit diseases accurately. The conventional diagnosis method of the diseases mainly relies on naked-eye symptom observations conducted by experienced agricultural technicians. This method is time-consuming and laboursome. In addition, there are not enough agricultural technicians to meet the actual needs of apple production. Therefore, it is necessary to explore a rapid, accurate, convenient, and highly automated disease identification method.

Image processing technology has been widely applied in the diagnosis, identification, and monitoring of plant diseases (Sankaran et al., 2010; Barbedo, 2016; Vishnoi et al., 2021), such as wheat diseases (Li et al., 2012; Johannes et al., 2017; Deng et al., 2021), maize diseases (DeChant et al., 2017; Chen et al., 2021), rice diseases (Phadikar et al., 2013; Lu et al., 2017; Narmadha et al., 2022), cotton diseases (Camargo and Smith, 2009; Caldeira et al., 2021), soybean diseases (Pires et al., 2016; Shrivastava et al., 2017; Araujo and Peixoto, 2019), cucumber diseases (Vakilian and Massah, 2013; Zhang S. W. et al., 2017; Kainat et al., 2021), tomato diseases (Yamamoto et al., 2017; Trivedi et al., 2021), grape diseases (Tian et al., 2007; Oberti et al., 2014; Zhu et al., 2020), and citrus diseases (Pydipati et al., 2006; Sankaran et al., 2013). Moreover, image processing technology has been used to make disease severity assessments (Li et al., 2011; Barbedo, 2014; Vieira et al., 2014; Shrivastava et al., 2015; Ganthaler et al., 2018), conduct pathogen identification (Chesmore et al., 2003; Deng et al., 2012; Wang et al., 2021), and perform automatic counting of pathogen spores (Li X. L. et al., 2013; Li et al., 2017). It is convenient and rapid to perform plant disease identification using image processing technology, and automatic disease identification can be realized, indicating that the image-based plant disease identification method has a good application prospect. However, most of the reported related studies focused on the diagnosis and identification of plant leaf diseases.

There have been some reports on image-based distinction and recognition of apple diseases (Yin et al., 2012; Huo et al., 2013; Dubey and Jalal, 2014, 2016; Omrani et al., 2014; Tan et al., 2015; Wang et al., 2015; Zhang C. L. et al., 2017; Liu et al., 2018; Bansal et al., 2021; Ortega-Sánchez et al., 2022), but few of them focused on the distinction and recognition of apple fruit diseases. The two related studies conducted by Yin et al. (2012) and Huo et al. (2013), respectively, were based on 78 low-resolution images of three kinds of apple fruit diseases including apple ring rot, apple anthracnose, and new apple ring rot (26 images per apple fruit disease) that were taken by using mobile phone in natural scenes, an improved level set interactive segmentation method was used to perform segmentation operation on the preprocessed images, and then six color features, eight texture features, and seven shape features were extracted. Based on the 15 texture and shape features, Yin et al. (2012) developed a support vector machine (SVM) model with a linear kernel function to identify the three kinds of apple fruit diseases, and average identification accuracy of 90.00% was achieved. Based on the eight texture features extracted from the segmented disease images, Huo et al. (2013) built the identification models of the three kinds of apple fruit diseases using three methods including gray relation analysis, SVM, and compressive sensing, the average identification accuracies of 86.67, 90, and 90%, respectively, were obtained for the three models, respectively. Tan et al. (2015) used a deep learning neural network based on flexible momentum to identify the images of diseased apple fruits and achieved a recall rate of 98.4%. Wang et al. (2015) developed a convolutional neural network (CNN) based on a variable impulse learning algorithm to conduct the identification of 100 images of diseased apple fruits, and the identification accuracy was 97.45%. The overall accuracy of 91.1% was obtained by Nachtigall et al. (2017) using CNN to identify healthy apple fruits and unhealthy apple fruits in five disorders including scab caused by *Venturia inaequalis*, alternaria rot caused by *Alternaria alternata*, bull's eye rot caused by *Cryptosporiopsis perennans*, penicillium rot caused by *Penicillium expansum*, and bitter pit (calcium deficiency) based on the images acquired under controlled conditions.

The quality of plant disease images acquired in natural scenes is usually affected by many factors such as uneven illumination, complex background, and blurred edges. It is necessary to explore an accurate and highly automated image segmentation method to segment these disease images. Furthermore, feature extraction after image segmentation is particularly important for image recognition. The local binary pattern (LBP) operator is a local texture descriptor (Ojala et al., 1996). Because of its characteristics of gray-scale invariance, simple calculation, and insensitivity to illumination changes, this operator is widely used in the fields such as medical image recognition (Nanni et al., 2012; Panda et al., 2018) and face recognition (Ahonen et al., 2006; Yang and Chen, 2013; Lu et al., 2018). The LBP operator has also been applied in the image-based recognition of plant diseases. Leiva-Valenzuela and Aguilera (2013) implemented image-based detection of fungally decayed, shriveled, and mechanically damaged blueberries based on the 951 extracted features including LBP features. In a study conducted by Dubey and Jalal (2014), based on apple fruit images of apple blotch, apple rot, apple scab, and normal apple, the K-means clustering technique was applied to perform image segmentation, color and texture features including global color histogram, color coherence vector, color difference histogram, structure element histogram, local ternary pattern, completed local binary pattern (CLBP), and LBP was extracted, and then the color, texture, and fused features were applied to identify apple fruit images by using a multi-class support vector machine (MSVM), finally, an average identification accuracy of approximately 90% was obtained. In another study conducted by Dubey and Jalal (2016), the MSVM method was used to classify apple blotch, apple rot, and apple scab

based on the color (global color histogram and color coherence vector), texture (LBP and CLBP), and shape (Zernike moments) feature extracted from apple fruit images, and the results showed that classification performance with accuracy more than 90% could be achieved by using the MSVM models built based on CLBP or each feature combination containing color and texture features. Multi-scale block local binary pattern (MB-LBP), a modified LBP operator, can extract texture information at different scales of an image and is not easily affected by image noise (Liao et al., 2007; Zhang et al., 2007). It has been applied in studies on object detection and recognition (Halidou et al., 2014; Li et al., 2015; Kang et al., 2017; Karanwal, 2021). To the best of our knowledge, there are no reports on the distinction between apple ring rot and apple anthracnose by using the image processing method based on LBP features.

In this study, after preprocessing the digital images of apple fruits infected with ring rot and anthracnose acquired in natural scenes, two lesion segmentation methods based on circle fitting were developed and applied to implement lesion segmentation of the disease images. Subsequently, the gray images on the nine components in the RGB, HSI, and L*a*b* color spaces of the segmented lesion images were filtered by using MB-LBP with pixel blocks in different sizes, and the corresponding LBP histogram features were extracted. Finally, based on these features, SVM models and random forest models were developed to distinguish the two kinds of apple fruit diseases. The aim of this study was to provide a rapid and accurate method for the non-destructive distinction of the two diseases in apple fruits.

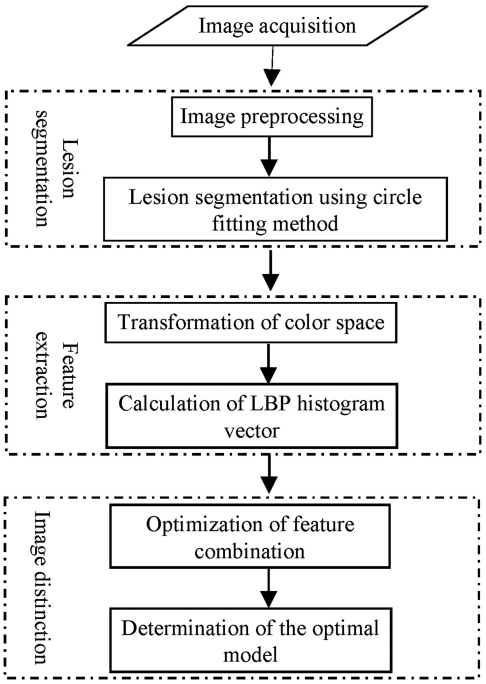


FIGURE 1 | Flow chart of image-based distinction of ring rot and anthracnose on apple fruits.

MATERIALS AND METHODS

The distinction of ring rot and anthracnose on apple fruits based on image processing was conducted in accordance with the procedures as shown in **Figure 1**.

Acquisition of Disease Images

A total of 123 apple fruit disease images were acquired using different digital cameras under field conditions, including 60 images of apple ring rot and 63 images of apple anthracnose. Most of the disease images used in this study were acquired in the apple orchards in Shangzhuang Experimental Station of China Agricultural University and Sujiatuo County, Haidian District, Beijing, China in the autumn of 2014 by using two digital cameras Canon PowerShot SX100 IS (Canon Inc., Tokyo, Japan) and Canon EOS 700D (Canon Inc., Tokyo, Japan), and were acquired in the apple orchards in Sujiatuo County, Haidian District, Beijing, China in the autumn of 2015 by using the digital camera Canon PowerShot SX100 IS. The other disease images were provided by Shutong Wang from the College of Plant Protection, Hebei Agricultural University, Baoding, China, and He Wang from Forest Pest Management and Quarantine Station of Beijing, Beijing, China.

Image Preprocessing

Preprocessing operations of the acquired disease images, including image scaling, color image contrast stretching, and morphological opening and closing reconstruction were performed by using the image processing toolbox in the software MATLAB R2013b (MathWorks, Natick, MA, United States).

Image Scaling

Because the disease images were acquired by using different digital cameras with different settings, there were obvious differences among the images in size. In order to process the disease images by using the same morphological operations, it is necessary to resize them into the same size range. In this study, the acquired disease images were scaled with an equal ratio in the range of 1,000 × 1,000 pixels to 2,000 × 2,000 pixels.

Color Image Contrast Stretching

Color image contrast stretching is conducive to the enhancement of the color difference between the lesions and the surrounding background, and this operation facilitates the subsequent lesion image segmentation. Color image contrast stretching was operated by using the following MATLAB code: *rgbstr* = imadjust(*rgb*, stretchlim(*rgb*)), where *rgb* is the RGB color image to be processed, and *rgbstr* is the processed image.

Morphological Opening and Closing by Reconstruction Operations

Morphological opening and closing by reconstruction operations can reduce the noise interference to the real edges (Wang et al., 2008; Zhang and Wang, 2009). In this study, the morphological opening and closing by reconstruction operations of *R*, *G*, and *B* color components were performed using the circular structure element (disk) with a radius of 10, and then the obtained three

color component images were integrated into a new color image by using the function "cat" in the MATLAB R2013b software.

Lesion Image Segmentation

The backgrounds of apple disease images obtained in natural scenes are usually complex, mainly including soil, branches, and leaves of apple trees, other green plants, and local strong reflection. Automatic and accurate segmentation of lesions on the surface of apple fruit from the complex backgrounds is crucial for disease distinction and identification. Generally, the lesions of ring rot and anthracnose on apple fruits have two distinct characteristics. Firstly, the lesions are usually brown, thus the red component of the lesions in a disease image is often greater than the green component. This characteristic can be used to distinguish green elements such as the leaves of apple trees and other green plants in the background. Secondly, the lesions are located on the surface of apple fruit and are usually nearly round, and this characteristic can be used to distinguish branches of apple trees, soil, and other backgrounds with similar colors to the lesions. In this study, according to the above characteristics, the approximate position of apple fruits in a disease image was determined firstly, and then lesion segmentation was conducted by using the lesion image segmentation methods based on circle fitting. The steps for lesion image segmentation in detail are as follows.

Step 1. To make full use of the image color information, the gradient of the integrated color image obtained after morphological opening and closing by reconstruction operations was first calculated. Assuming that $c(x, y)$ is the gradient of any point (x, y) in the color image, it can be calculated according to the method described by Gonzalez and Woods (2011), which can be expressed as follows.

Let r, g, and b be the unit vectors of the R-axis, G-axis, and B-axis in the RGB color space, respectively, and the vectors u and v can be defined as:

$$u = \frac{\partial R}{\partial x} r + \frac{\partial G}{\partial x} g + \frac{\partial B}{\partial x} b \tag{1}$$

and

$$v = \frac{\partial R}{\partial y} r + \frac{\partial G}{\partial y} g + \frac{\partial B}{\partial y} b \tag{2}$$

Let g_{xx}, g_{yy}, and g_{xy} represent the dot products of these vectors u and v, as follows:

$$g_{xx} = u \cdot u = u^T u = \left|\frac{\partial R}{\partial x}\right|^2 + \left|\frac{\partial G}{\partial x}\right|^2 + \left|\frac{\partial B}{\partial x}\right|^2 \tag{3}$$

$$g_{yy} = v \cdot v = v^T v = \left|\frac{\partial R}{\partial y}\right|^2 + \left|\frac{\partial G}{\partial y}\right|^2 + \left|\frac{\partial B}{\partial y}\right|^2 \tag{4}$$

and

$$g_{xv} = u \cdot v = u^T v = \frac{\partial R}{\partial x}\frac{\partial R}{\partial y} + \frac{\partial G}{\partial x}\frac{\partial G}{\partial y} + \frac{\partial B}{\partial x}\frac{\partial B}{\partial y} \tag{5}$$

here, the direction of the maximum change rate of $c(x, y)$ can be given by the angle $\theta(x, y)$, which can be calculated by using the following formula:

$$\theta(x, y) = \frac{1}{2}\arctan\left[\frac{2g_{xy}}{g_{xx} - g_{yy}}\right] \tag{6}$$

and the value of the change rate at point (x, y) in the direction of the angle $\theta(x, y)$ can be given by using the following formula:

$$F_\theta(x, y) = \left\{\frac{1}{2}\left[(g_{xx} + g_{xy}) + (g_{xx} - g_{xy})\cos 2\theta(x, y) + 2g_{xy}\sin 2\theta(x, y)\right]\right\}^{\frac{1}{2}} \tag{7}$$

The partial derivatives of Formulas (3), (4), and (5) can be calculated by using the Sobel operator, and then the gradient of any point (x, y) can be calculated.

Step 2. Edge detection of the gradient image generated in Step 1 was carried out by using the Canny operator. For the Canny operator, the default values were used for the sensitivity thresholds, and the standard deviation of the Gaussian filter, σ(sigma), was set to 20. The purpose of edge detection is to preserve the real edge, remove the false edge, and present the edge image in a binary pattern. In the binary edge image, the edge of the junction of the diseased and healthy regions may have breakpoints. Therefore, in this study, the circular structure element with a radius of 2 was used to conduct a dilation operation on the edge image to obtain the continuous lesion edge as much as possible.

Step 3. To remove the green background such as the leaves of the apple tree and other green plants from the image, the pixel points with the green component greater than the red component were assigned a value of 0, and the other pixel points were assigned a value of 1, and then the green-background-subtracted binary image can be obtained. The green apple fruits in the image can also be removed in the process of the background subtraction. Because the lesions of the two apple fruit diseases are usually brown for which the green component is much smaller than the red component, they could still be completely retained in the image after the background subtraction, and thus the subsequent lesion extraction will not be affected. The binary edge image obtained in Step 2 was inversed, and then multiplied with the green-background-subtracted binary image, thus a new binary image was obtained.

Step 4. Cavity filling of the binary image obtained in Step 3 was carried out. To avoid the adhesion between apple fruits and the background in the image and remove the relatively small background target, the circular structure element with a radius of 50 was used to conduct the opening operation on the binary image obtained after cavity filling.

Step 5. The areas of all the remaining connected components were calculated, and any region for which the area of the connected component was less than two-thirds of the area of the maximum connected component was removed. The retained regions were regarded as the regions where apple fruits may exist.

Step 6. Assuming that the number of the retained regions was M, that is, the number of the regions where apple fruits may exist was M, let $j = 1$, then the convex hull of the jth apple fruit region was calculated and the region contour was extracted. Considering

that apple fruits are usually near-circular, circle fitting of the *j*th apple fruit region was carried out based on the contour combined with the least square method. Circular curve fitting by using the least square method was carried out according to the specific calculation method as described by Li and He (2013), which can be listed as follows.

Suppose the formula of the circular curve to be fitted is:

$$Rad^2 = (x - Ax)^2 + (y - By)^2 \tag{8}$$

in which *Rad* is the radius of the circle curve, and *Ax* and *By* are the abscissa and ordinate of the circle center, respectively.

Suppose the pixel point set on the contour line is (x_i, y_i) where $i = 1, 2, \ldots, N$, and N represent the number of pixel points on the contour line. The distance from the *i*th pixel point in the set to the circle center is d_i, then

$$d_i^2 = (x_i - Ax)^2 + (y_i - By)^2 \tag{9}$$

The difference between the square of the distance from the point (x_i, y_i) to the circle center and the square of the radius of the circular curve, δ_i, can be described as the following formula:

$$\begin{aligned}\delta_i = d_i^2 - Rad^2 &= (x_i - Ax)^2 + (y_i - By)^2 - Rad^2 \\ &= x_i^2 + y_i^2 + ax_i + by_i + c\end{aligned} \tag{10}$$

Let $Q(a, b, c)$ be the sum of squares of δ_i, then

$$Q(a, b, c) = \sum \delta_i^2 = \sum \left[x_i^2 + y_i^2 + ax_i + by_i + c\right]^2 \tag{11}$$

The least-square method was used to calculate and achieve the optimal fitted circular curve, that is, the parameters *a*, *b*, and *c* were calculated to minimize $Q(a, b, c)$. The partial derivatives of $Q(a, b, c)$ with respect to *a*, *b*, and *c* were calculated, respectively, and then were set to 0, thus the extreme points could be achieved and the values of the corresponding parameters *a*, *b*, and *c* could be obtained. The partial derivatives of $Q(a, b, c)$ with respect to *a*, *b*, and *c* were calculated according to the following Formulas (12), (13), and (14), respectively.

$$\frac{\partial Q(a, b, c)}{\partial a} = \sum 2\left(x_i^2 + y_i^2 + ax_i + by_i + c\right)x_i = 0 \tag{12}$$

$$\frac{\partial Q(a, b, c)}{\partial b} = \sum 2\left(x_i^2 + y_i^2 + ax_i + by_i + c\right)y_i = 0 \tag{13}$$

$$\frac{\partial Q(a, b, c)}{\partial c} = \sum 2\left(x_i^2 + y_i^2 + ax_i + by_i + c\right) = 0 \tag{14}$$

After the values of the corresponding parameters *a*, *b*, and *c* were obtained, the fitted values of *Ax*, *By*, and *Rad* could be estimated according to the following Formulas (15), (16), and (17), respectively.

$$Ax = \frac{a}{-2} \tag{15}$$

$$By = \frac{b}{-2} \tag{16}$$

$$Rad = \frac{1}{2}\sqrt{a^2 + b^2 - 4c} \tag{17}$$

After circle fitting of the *j*th apple fruit contour was conducted, the distances between all the pixel points in the image and the circle center were calculated, then the pixel points with a distance less than the circle radius were assigned a value of 1 and the other pixel points were assigned a value of 0, and thus the binary image of the *j*th apple fruit region was obtained.

Step 7. Two methods were tried to find the real lesion edges in this study, thus the corresponding lesion image segmentation methods were classified as Lesion segmentation method 1 and Lesion segmentation method 2, respectively. For Lesion segmentation method 1, the binary edge image obtained in Step 2 was multiplied with the circle fitting binary image of the *j*th apple fruit region, in order to retain the edge inside the apple fruit region and remove the edge outside the apple fruit region; the convex hull areas of all the edges were calculated, and the edge with the largest convex hull area was considered as the edge of the junction of the diseased and healthy regions; then the convex hull contour corresponding to this edge was calculated, and the circle fitting method of apple fruit region was used to fit the contour, finally the region obtained by circle fitting was treated as the region where the disease lesion was located. For Lesion segmentation method 2, firstly, the circular structure element with a radius of 20 was used to conduct the closing operation on the image obtained by the opening operation in Step 4, and a binary image was obtained. The purpose of this operation was to reduce the possible depressions in the retained regions. Subsequently, the binary edge image obtained in Step 2 was multiplied with the circle fitting binary image of the *j*th apple fruit region and then was multiplied with the above binary image obtained by closing operation, aiming to retain the edge inside the apple fruit region and remove the edge outside the apple fruit region. The remaining procedures of Lesion segmentation method 2 were the same as that of Lesion segmentation method 1.

Step 8. Let $j = j+1$, if $j \leq M$, then Steps 6 and 7 will be repeated, otherwise the operations for the lesion image segmentation will be finished.

After segmentation, each pixel in a lesion image was determined as a lesion pixel or a healthy pixel. The evaluation of image segmentation performance can be conducted by referring to the evaluation method of a binary classification model (Powers, 2011). The manual segmentation method using the Adobe Photoshop CC software was utilized to conduct segmentation of the lesion images, and the segmentation results were considered as references. The segmentation results obtained by using the manual segmentation method were compared with those obtained by using the two segmentation methods described above, and Recall, Precision, and Score (Qin et al., 2016) were used as the indices to evaluate the above two segmentation methods. The three indices were calculated according to the following formulas as described by Qin et al. (2016):

$$\text{Recall} = \frac{N_1}{N_2} \tag{18}$$

$$\text{Precision} = \frac{N_1}{N_3} \tag{19}$$

$$\text{Score} = \frac{\text{Recall} + \text{Precision}}{2} \tag{20}$$

where N_1 is the total number of lesion pixels in a lesion image correctly determined by using one of the segmentation methods described above, N_2 is the total number of lesion pixels in the lesion image determined by using the manual segmentation method, and N_3 is the total number of the pixels in the lesion image. All of these three indices range from 0 to 1. To evaluate the performances of the two lesion segmentation methods, the image dataset of apple ring rot comprising 60 images, the image dataset of apple anthracnose comprising 63 images, and the aggregated image dataset comprising all of the 123 images, were constructed by using the acquired images after preprocessing. The segmentation method with larger values of Recall, Precision, and Score, was chosen as the automatic lesion segmentation method for further disease image processing and disease distinction.

Extraction of Local Binary Pattern Histogram Features From the Segmented Lesion Images

To reduce the influence of illumination on the image features, MB-LBP operators with pixel blocks in different sizes were used to filter the gray images on the nine components in the RGB, HSI, and L*a*b* color spaces of the segmented lesion images, and then the corresponding LBP histogram vectors were calculated.

Initially, all of the segmented lesion images were resized to 256 × 256 pixels. The three-scale sizes of pixel blocks (sub-regions) were set as 1 × 1 pixels, 2 × 2 pixels, and 3 × 3 pixels, and the corresponding MB-LBP operators were recorded as MB_1-LBP, MB_2-LBP, and MB_3-LBP, respectively. For each MB-LBP operator, the number of neighborhoods was set to 8, and the neighborhood radius was set to 2. The MB_3-LBP operator, as shown in **Figure 2**, was taken as an example. In **Figure 2**, each small square surrounded by thin black lines represents a pixel, and each square (sub-region) enclosed by thick black lines represents a pixel block. The gray value of each pixel block is the average of the gray values of nine (3 × 3) pixels included in the corresponding block. The black point at the center of the center block is labeled as the center point, and the eight surrounding black points are labeled as to its eight neighborhood points. For the black point located at the center of a pixel block, the gray value of the pixel block is used directly as the value of this black point. For the black point is not located at the center of a pixel block, the gray value of the pixel block is determined by using the bilinear interpolation method. By comparing the gray value of each neighborhood point with that of the center point, an 8-bit binary number is obtained, which is then used as the response value of the center point. MB_1-LBP operator and MB_2-LBP operator are the same as the MB_3-LBP operator except for the size of each pixel block.

The gray images on the nine components in the RGB, HSI, and L*a*b* color spaces of the segmented lesion images were filtered by using MB_1-LBP with the size of the pixel block of 1 × 1, MB_2-LBP with the size of the pixel block of 2 × 2, and MB_3-LBP with the size of the pixel block of 3 × 3, respectively, and the corresponding local binary pattern histogram vectors were calculated as the features of the lesion images (Zhang et al., 2013). In this study, the uniform LBP operator with 59 histogram bins that include 58 uniform histogram bins and one non-uniform histogram bin, was used for the calculation of the LBP histogram. Finally, the LBP histogram vector in 59 dimensions was obtained. The algorithm in detail was described by Zhang et al. (2007, 2013).

Disease Distinction Model Building Based on Local Binary Pattern Histogram Features of the Segmented Lesion Images

From all of the acquired apple disease images, 40 images of ring rot and 42 images of anthracnose were randomly selected to form the training set, and the remaining 20 images of ring rot and the remaining 21 images of anthracnose were used to form the testing set. Disease distinction models were built by using two modeling methods including the SVM method and the random forest method.

The LBP histogram features extracted by using the MB-LBP operators are in a large number of dimensions. The SVM method can be applied to effectively solve the data problems of small samples, non-linearity, high dimensions, and local minima (Cortes and Vapnik, 1995; Burges, 1998). Therefore, the SVM method was used to build distinction models of the images of ring rot and anthracnose on apple fruits in this study. Based on the LBP histogram features extracted from the segmented lesion images, the SVM models for the distinction of the two kinds of apple diseases were built by using C-SVM in the LIBSVM package developed by Chang and Lin (2011). To build an SVM model, a radial basis function kernel was selected, and the grid search algorithm was used to search for the optimal penalty parameter C and the optimal kernel function parameter g in the range of 2^{-10}–2^{10} with a searching step of 0.4. Based on the training set, the distinction accuracies at all points within the grid were achieved by running three complete cross-validations. When the highest distinction accuracy was achieved, the corresponding values of C and g were treated as the optimal parameters and were recorded as C_{best} and g_{best}, respectively. Then, the SVM model was built by using the parameters C_{best} and g_{best}. The distinction accuracies of the training set and testing set were calculated and were used to evaluate the model distinction performance.

A random forest, composed of multiple decision trees, can realize prediction by integrating the prediction result of each decision tree (Breiman, 2001). This modeling method can deal well with high-dimensional features, and the running speed of the built model is fast. Therefore, the random forest method was used to build distinction models of the images of apple ring rot and apple anthracnose based on the extracted LBP histogram features in this study. To a certain extent, the distinction performance of a random forest model depends on the number of decision trees constituting the model, so it is necessary to try a variety

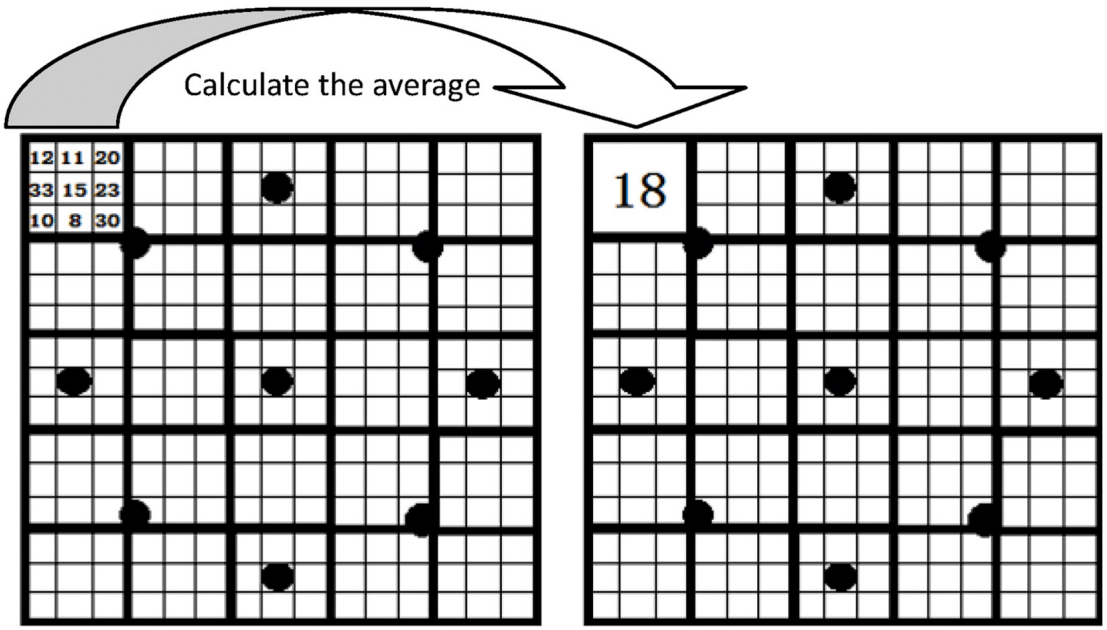


FIGURE 2 | The diagram of the MB_3-LBP operator.

of values and determine the optimal number of decision trees according to the distinction performances of the built random forest models. In this study, during building the random forest models for disease distinction, the number of decision trees was successively assigned as 10, 20, 30, 40, 50, 60, 70, 80, 90, and 100, and the optimal random forest model was determined according to the distinction accuracies of the training set and testing set and the number of decision trees. For the built random forest models with the same distinction accuracies of the training set and testing set, the one with the least number of decision trees was considered the optimal model. For each decision tree, the arithmetic square root ($\sqrt{N}$ or $N^{1/2}$) of the total number (N) of LBP histogram features used for modeling was treated as the number of features randomly selected. If $\sqrt{N}$ or $N^{1/2}$ was a decimal, the integral number obtained by rounding up the decimal was considered the value of the feature number.

RESULTS

Results of Image Preprocessing and Image Segmentation

For apple ring rot, the results of image preprocessing and image segmentation, taking an image as an example, are shown in **Figure 3**. As shown in **Figures 3A,B**, after image preprocessing, the color of the lesion region on the surface of the diseased apple was obviously deepened, and the edge of the junction of the diseased and healthy regions became clearer. The results of lesion image segmentation by using Lesion segmentation method 1 and Lesion segmentation method 2 are shown in **Figures 3C,D**, respectively. The results demonstrated that the size and location of the lesion segmented by using Lesion segmentation method 1 were closer to that of the real lesion than that of the lesion segmented by using Lesion segmentation method 2. For apple anthracnose, there were no relatively obvious differences between the segmentation results of the two segmentation methods, and the satisfactory lesion segmentation performances were achieved by using both Lesion segmentation method 1 and Lesion segmentation method 2. Taking an image of apple anthracnose as an example, the corresponding results of image preprocessing and image segmentation are shown in **Figure 4**.

After lesion segmentation operations of all the diseased images were conducted by using the two lesion segmentation methods (i.e., Lesion segmentation method 1 and Lesion segmentation method 2), the statistical results of Recalls, Precisions, and Scores for the two methods based on the three image datasets described above are shown in **Table 1**. The shape of very few lesions was very irregular, e.g., two or more lesions joined together, resulting in the extreme values of Recall, Precision, and Score. To reduce the influence of the extreme values on evaluating the lesion segmentation methods, the mean and median of each evaluation index (Recall, Precision, or Score) were used to evaluate the performances of the two lesion segmentation methods described above.

For the image dataset of apple ring rot, when Lesion segmentation method 1 was used, the mean and median of the Recalls were 0.93 and 0.99, respectively; the mean and median of the Precisions were 0.92 and 0.95, respectively; and the mean and median of the Scores were 0.93 and 0.95, respectively. For this image dataset, when Lesion segmentation method 2 was used, the mean and median of the Recalls were 0.81 and 0.89, respectively; the mean and median of the Precisions were 0.93 and 0.96, respectively; and the mean and median of the Scores were 0.87 and 0.93, respectively. The results demonstrated that,

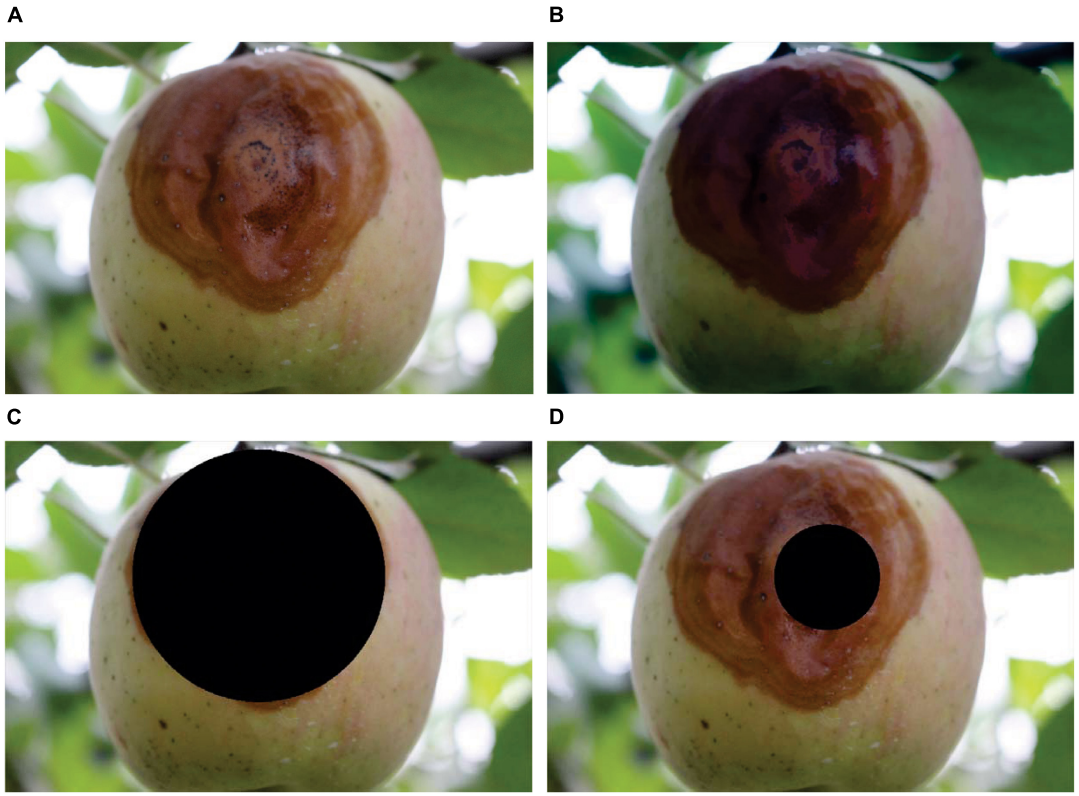


FIGURE 3 | The results of image preprocessing and lesion segmentation for apple ring rot. **(A)** Original color image. **(B)** Image after preprocessing. **(C)** Image after lesion segmentation by using Lesion segmentation method 1 without closing operation. **(D)** Image after lesion segmentation by using Lesion segmentation method 2 with the closing operation. The black areas in panels **(C,D)** are the segmented lesions obtained by using Lesion segmentation method 1 and Lesion segmentation method 2, respectively.

for the image dataset of apple ring rot, the means and medians of the Recalls and Scores obtained when Lesion segmentation method 1 was used, were higher than those obtained when Lesion segmentation method 2 was used; and the mean and median of the Precisions were similar when Lesion segmentation method 1 and Lesion segmentation method 2 were used, respectively. The results indicated that Lesion segmentation method 1 was more suitable for lesion segmentation of the images of ring rot on apple fruits.

For the image dataset of apple anthracnose, when Lesion segmentation method 1 was used, the mean and median of the Recalls were 0.95 and 1, respectively; the mean and median of the Precisions were 0.94 and 0.97, respectively; and the mean and median of the Scores were 0.94 and 0.97, respectively. When Lesion segmentation method 2 was used on this image dataset, the mean and median of the Recalls were 0.91 and 0.99, respectively; the mean and median of the Precisions were 0.96 and 0.97, respectively; and the mean and median of the Scores were 0.93 and 0.97, respectively. The results demonstrated that, for the image dataset of apple anthracnose, the mean and median of the Recalls, Precisions, or Scores obtained when Lesion segmentation method 1 and Lesion segmentation method 2 were used, respectively, were similar, and all the indices were more than 0.9. The results indicated that these two lesion segmentation methods were both suitable for lesion segmentation of the images of apple anthracnose.

For the image dataset of the aggregated image dataset obtained after aggregation of the two image datasets of apple ring rot and apple anthracnose, when Lesion segmentation method 1 was used, the mean and median of the Recalls were 0.94 and 0.99, respectively; the mean and median of the Precisions were 0.93 and 0.96, respectively; and the mean and median of the Scores were 0.93 and 0.96, respectively. For this aggregated image dataset, when Lesion segmentation method 2 was used, the mean and median of the Recalls were 0.86 and 0.96, respectively; the mean and median of the Precisions were 0.95 and 0.97, respectively; and the mean and median of the Scores were 0.9 and 0.95, respectively. The results showed that, for the aggregated image dataset, the mean and median of the Recalls or Scores obtained when Lesion segmentation method 1 was used, were both higher than those obtained when Lesion segmentation method 2 was used, and the mean and median of the Precisions obtained when the former method was used were similar to those obtained when the latter method was used.

The results described above indicated that Lesion segmentation method 1 was more suitable for lesion segmentation of the images of ring rot and anthracnose on apple fruits. Therefore, Lesion segmentation method 1 was

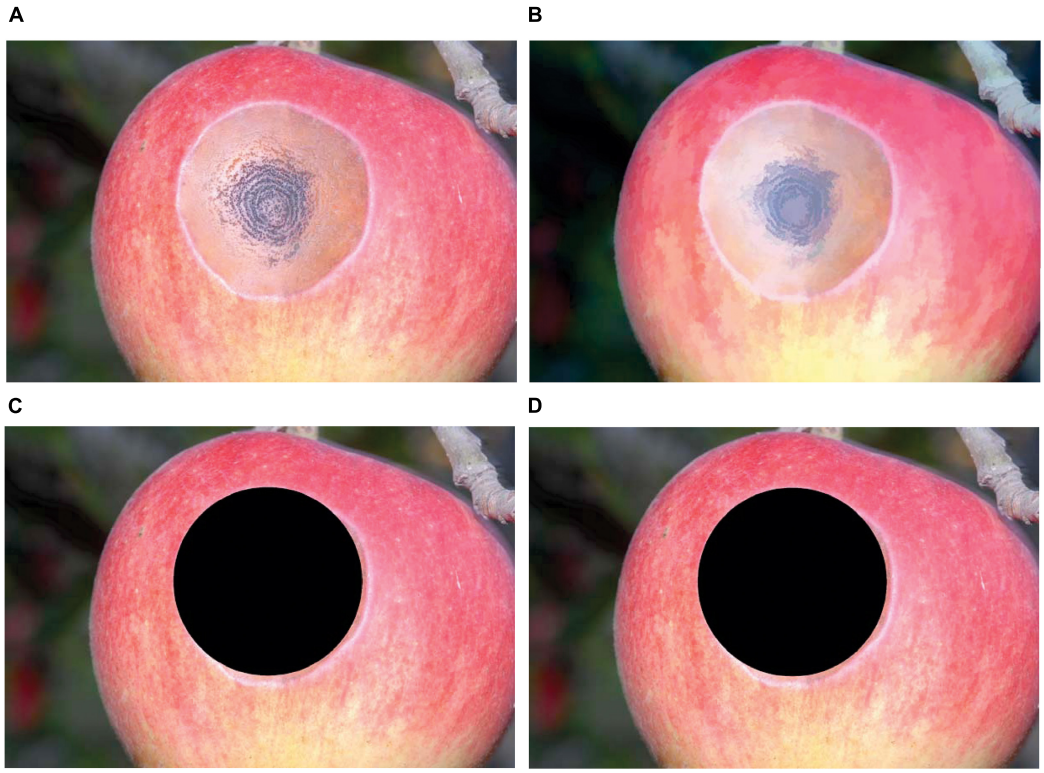


FIGURE 4 | The results of image preprocessing and lesion segmentation for apple anthracnose. **(A)** Original color image. **(B)** Image after preprocessing. **(C)** Image after lesion segmentation by using Lesion segmentation method 1 without closing operation. **(D)** Image after lesion segmentation by using Lesion segmentation method 2 with the closing operation. The black areas in panels **(C,D)** are the segmented lesions obtained by using Lesion segmentation method 1 and Lesion segmentation method 2, respectively.

selected for realizing the automatic segmentation of the lesion images of the two apple fruit diseases in this study.

Distinction Results of the Support Vector Machine Models Based on the Local Binary Pattern Histogram Features

The distinction results of the SVM models based on the LBP histogram features of the gray images on each individual component in the RGB, HSI, and L*a*b* color spaces of the segmented lesion images are shown in **Tables 2–4**, respectively. R1 denoted the LBP histogram feature of the gray image of the *R* component of the lesion image filtered by the MB_1-LBP operator, and R2 denoted the LBP histogram feature of the gray image of the *R* component of the lesion image filtered by the MB_2-LBP operator, R1G1 denoted the combination of R1 and G1, and the rest features' names could be deduced by analogy. The results showed that the optimal SVM model for disease distinction was built based on the feature L1a1 and that this SVM model had the best distinction performance. For this optimal SVM model, the parameters C_{best} and g_{best} were 12.126 and 0.144, respectively, and the distinction accuracy of the training set was

TABLE 1 | Statistical comparison of the segmentation effects using the two lesion segmentation methods.

Image dataset	Lesion segmentation method	Recall		Precision		Score	
		Mean	Median	Mean	Median	Mean	Median
Image dataset of apple ring rot	Lesion segmentation method 1	0.93	0.99	0.92	0.95	0.93	0.95
	Lesion segmentation method 2	0.81	0.89	0.93	0.96	0.87	0.93
Image dataset of apple anthracnose	Lesion segmentation method 1	0.95	1.00	0.94	0.97	0.94	0.97
	Lesion segmentation method 2	0.91	0.99	0.96	0.97	0.93	0.97
Aggregated image dataset	Lesion segmentation method 1	0.94	0.99	0.93	0.96	0.93	0.96
	Lesion segmentation method 2	0.86	0.96	0.95	0.97	0.90	0.95

Aggregated image dataset was obtained after aggregation of the two image datasets of apple ring rot and apple anthracnose.

TABLE 2 | Distinction results of the SVM models based on the LBP histogram features of the gray images of the three components in RGB color space of the segmented lesion images.

Feature	The optimal parameters of SVM model		Distinction accuracy of the training set/%	Distinction accuracy of the testing set/%
	C_{best}	g_{best}		
R1	21.112	0.082	93.90	90.24
G1	2.297	0.144	80.49	70.73
B1	0.758	0.250	78.05	60.98
R1G1	2.297	0.144	92.68	85.37
R1B1	21.112	0.047	96.34	92.68
G1B1	588.134	0.005	96.34	73.17
R1G1B1	194.012	0.016	100.00	87.80
R2	1.320	2.297	100.00	70.73
G2	2.297	0.435	96.34	58.54
B2	0.758	0.250	75.61	65.85
R2G2	111.430	0.005	95.12	80.49
R2B2	1.320	0.082	87.80	82.93
G2B2	0.758	0.144	78.05	68.29
R2G2B2	1.320	0.082	90.24	80.49
R3	2.297	2.297	100.00	73.17
G3	337.794	0.003	90.24	58.54
B3	21.112	0.027	90.24	68.29
R3G3	2.297	0.250	98.78	75.61
R3B3	21.112	0.027	96.34	82.93
G3B3	64	0.047	100.00	58.54
R3G3B3	4	0.082	97.56	75.61

R1 represents the LBP histogram feature of the gray image of the R component of the lesion image filtered by the MB_1-LBP operator; R2 represents the LBP histogram feature of the gray image of the R component of the lesion image filtered by the MB_2-LBP operator; R1G1 represents the combination of R1 and G1; the rest features' implication could be deduced by analogy.

100% and the distinction accuracy of the testing set was 95.12%. The model for which the distinction performance ranked second among all the built SVM models, was built based on the feature R1B1 with the parameters C_{best} and g_{best} of 21.112 and 0.047. For this model, the distinction accuracies of the training set and testing set were 96.34% and 92.68%, respectively. The model for which the distinction performance ranked third among all the built SVM models, was built based on the feature R1 with the parameters C_{best} and g_{best} of 21.112 and 0.082. For this model, the distinction accuracy of the training set was 93.9% and the distinction accuracy of the testing set was 90.24%. The SVM model was built based on the feature R1G1B1 with the optimal parameters C_{best} and g_{best} of 194.012 and 0.016 and the SVM model was built based on the feature L1a1b1 with the optimal parameters C_{best} and g_{best} of 6.964 and 0.082, the distinction accuracies of the training set were both 100% and the distinction accuracies of the testing set were both 87.8%. For the SVM model built based on the feature L1b1 with the optimal parameters, C_{best} and g_{best} of 2.297 and 0.25, the distinction accuracies of the training set and testing set were 97.56 and 85.37%, respectively. For the SVM model built based on the feature R1G1 with the optimal parameters, C_{best} and g_{best} of 2.297 and 0.144, the distinction accuracies of the training set and testing set were 92.68 and 85.37%, respectively. The results demonstrated that accurate distinction of apple ring rot and apple anthracnose can be achieved by using the SVM modeling method based on LBP histogram features. The LBP histogram features used in the above SVM models with satisfactory distinction performances were obtained by filtering the gray images of the related components with the MB_1-LBP operator. Compared with the other two MB-LBP operators (MB_2-LBP and MB_2-LBP), the pixel block of the MB_1-LBP operator is the smallest, and the highest fineness of the image texture can be obtained after image filtering with it, which may be helpful to improve the ability of the models to distinguish between apple ring rot and apple anthracnose.

Distinction Results of the Random Forest Models Based on the Local Binary Pattern Histogram Features

The distinction results of the random forest models built based on the LBP histogram features of the gray images of the three components in the RGB, HIS, and L*a*b* color spaces of the segmented lesion images of apple ring rot and apple anthracnose

TABLE 3 | Distinction results of the SVM models based on the LBP histogram features of the gray images of the three components in the HSI color space of the segmented lesion images.

Feature	The optimal parameters of SVM model		Distinction accuracy of the training set/%	Distinction accuracy of the testing set/%
	C_{best}	g_{best}		
H1	0.758	0.082	69.51	68.29
S1	111.430	0.016	82.93	82.93
I1	2.297	0.758	96.34	78.05
H1S1	6.964	0.016	75.61	82.93
H1I1	2.297	0.047	79.27	78.05
S1I1	0.435	0.144	78.05	78.05
H1S1I1	0.435	0.082	75.61	80.49
H2	2.297	0.144	78.05	75.61
S2	337.794	0.005	89.02	46.34
I2	0.758	0.250	79.27	68.29
H2S2	4.000	0.250	100.00	70.73
H2I2	0.758	0.144	81.71	80.49
S2I2	1.320	0.082	81.71	75.61
H2S2I2	0.758	0.250	92.68	80.49
H3	6.964	0.027	75.61	70.73
S3	0.758	0.758	93.90	63.41
I3	21.112	0.002	62.20	65.85
H3S3	1.320	0.144	90.24	75.61
H3I3	1.320	0.435	98.78	82.93
S3I3	1.320	0.758	100.00	70.73
H3S3I3	1.320	0.435	100.00	78.05

H1 represents the LBP histogram feature of the gray image of the H component of the lesion image filtered by the MB_1-LBP operator; H2 represents the LBP histogram feature of the gray image of the H component of the lesion image filtered by the MB_2-LBP operator; H1S1 represents the combination of H1 and S1; the rest features' implication could be deduced by analogy.

TABLE 4 | Distinction results of the SVM models based on the LBP histogram features of the gray images of the three components in the L*a*b* color space of the segmented lesion images.

Feature	The optimal parameters of SVM model		Distinction accuracy of the training set/%	Distinction accuracy of the testing set/%
	C_{best}	g_{best}		
L1	12.126	0.435	100.00	73.17
a1	2.297	1.320	100.00	65.854
b1	1.320	0.758	96.34	80.49
L1a1	12.126	0.144	100.00	95.12
L1b1	2.297	0.250	97.56	85.37
a1b1	2.297	0.758	100.00	70.73
L1a1b1	6.964	0.082	100.00	87.80
L2	1.320	0.758	95.12	63.41
a2	36.758	0.009	80.49	70.73
b2	1.320	0.144	84.15	75.61
L2a2	0.758	0.435	96.34	73.17
L2b2	0.758	0.250	90.24	73.17
a2b2	36.758	0.027	98.78	65.85
L2a2b2	1.320	0.082	95.12	78.05
L3	111.430	0.047	98.78	58.54
a3	337.794	0.009	91.46	56.10
b3	6.964	0.082	91.46	75.61
L3a3	1.320	0.435	100.00	70.73
L3b3	1.320	0.435	100.00	80.49
a3b3	0.758	0.144	87.80	78.05
L3a3b3	1.320	0.144	98.78	80.49

*L1 represents the LBP histogram feature of the gray image of the L * component of the lesion image filtered by the MB_1-LBP operator; L2 represents the LBP histogram feature of the gray image of the L * component of the lesion image filtered by the MB_2-LBP operator; L1a1 represents the combination of L1 and a1; the rest features' implication could be deduced by analogy.*

are shown in **Tables 5**–**7**, respectively. The results showed that the optimal random forest model for the distinction of the two apple fruit diseases was built with the number of decision trees equal to 30 based on the feature R1B1 and that the distinction performance of this model was the best among all the built random forest models. For this optimal model, the distinction accuracies of the training set and testing set were 100 and 90.24%, respectively. In terms of disease distinction performance, three models tied for second place with the distinction accuracies of the training set and testing set equal to 100 and 87.8%, respectively, among all the built random forest models. Among these three models, one was built with the number of decision trees equal to 90 based on the feature R1, another was built with the number of decision trees equal to 50 based on the feature R1G1B1, and the other was built with the number of decision trees equal to 80 based on the feature L3a3b3. For the random forest model built based on the feature R1G1, H1S1, H1S1I1, I3, H3I3, L1b1, or L3a3 with the number of decision trees corresponding to 60, 60, 90, 80, 100, 70, or 30, the distinction accuracy of the training set was 100% and the distinction accuracy of the testing set was 85.37%. The results demonstrated that accurate distinction of apple ring rot and apple anthracnose can be obtained by using the random forest method based on LBP histogram features.

DISCUSSION AND CONCLUSION

Diseases play important roles in the reduction of the yield and quality of apple fruits. Accurate disease diagnosis is a key prerequisite for the prevention and control of apple diseases. In this study, taking ring rot and anthracnose on apple fruits as the research objects, the lesion image segmentation of the two apple diseases was carried out, and then the MB-LBP features were extracted from the segmented lesion images, finally, the distinction of the images of the two apple diseases was conducted by using both the SVM method and the random forest method. According to the characteristics of the lesions of the two apple diseases, two lesion segmentation methods based on circle fitting, Lesion segmentation method 1 and Lesion segmentation method 2, were developed and compared. The statistical results of three evaluation indices including Recall, Precision, and Score of the two lesion segmentation methods indicated that Lesion segmentation method 1 was better than Lesion segmentation method 2. Therefore, Lesion segmentation method 1 was selected

TABLE 5 | Distinction results of the random forest models based on the LBP histogram features of the gray images of the three components in RGB color space of the segmented lesion images.

Feature	The number of decision trees built by the best random forest model	Distinction accuracy of the training set/%	Distinction accuracy of the testing set/%
R1	90	100.00	87.80
G1	60	100.00	75.61
B1	100	100.00	75.61
R1G1	60	100.00	85.37
R1B1	30	100.00	90.24
G1B1	80	100.00	80.49
R1G1B1	50	100.00	87.80
R2	90	100.00	78.05
G2	10	98.78	75.61
B2	20	100.00	63.41
R2G2	20	98.78	73.17
R2B2	60	100.00	82.93
G2B2	10	97.56	73.17
R2G2B2	90	100.00	82.93
R3	100	100.00	80.49
G3	70	100.00	73.17
B3	50	100.00	73.17
R3G3	60	100.00	80.49
R3B3	90	100.00	75.61
G3B3	90	100.00	73.17
R3G3B3	50	100.00	75.61

R1 represents the LBP histogram feature of the gray image of the R component of the lesion image filtered by the MB_1-LBP operator; R2 represents the LBP histogram feature of the gray image of the R component of the lesion image filtered by the MB_2-LBP operator; R1G1 represents the combination of R1 and G1; the rest features' implication could be deduced by analogy.

TABLE 6 | Distinction results of the random forest models based on the LBP histogram features of the gray images of the three components in the HSI color space of the segmented lesion images.

Feature	The number of decision trees built by the best random forest model	Distinction accuracy of the training set/%	Distinction accuracy of the testing set/%
H1	90	100.00	70.73
S1	30	100.00	78.05
I1	70	100.00	78.05
H1S1	60	100.00	85.37
H1I1	50	100.00	80.49
S1I1	50	100.00	82.93
H1S1I1	90	100.00	85.37
H2	20	100.00	75.61
S2	40	100.00	82.93
I2	40	100.00	73.17
H2S2	90	100.00	78.05
H2I2	60	100.00	82.93
S2I2	60	100.00	75.61
H2S2I2	90	100.00	80.49
H3	70	100.00	73.17
S3	20	100.00	70.73
I3	80	100.00	85.37
H3S3	60	100.00	80.49
H3I3	100	100.00	85.37
S3I3	50	100.00	80.49
H3S3I3	70	100.00	82.93

H1 represents the LBP histogram feature of the gray image of the H component of the lesion image filtered by the MB_1-LBP operator; H2 represents the LBP histogram feature of the gray image of the H component of the lesion image filtered by the MB_2-LBP operator; H1S1 represents the combination of H1 and S1; the rest features' implication could be deduced by analogy.

to realize automatic segmentation of the lesion images in this study. To reduce the influence of illumination on the features of lesion images, the gray images of the nine components in the RGB, HSI, and L*a*b* color spaces of the segmented lesion images of the two apple diseases were filtered by using MB_1-LBP, MB_2-LBP, and MB_3-LBP operators, respectively, and the corresponding LBP histogram features were extracted for further disease distinction. The results demonstrated that for the built disease distinction SVM model based on the feature R1, R1G1, R1B1, L1a1, L1b1, R1G1B1, or L1a1b1, the distinction accuracies of the training set and testing set were high, and the satisfactory disease distinction performance was achieved. Among these SVM models, the distinction performance of the model built based on the feature L1a1 was optimal. The obtained results showed that the satisfactory disease distinction performance could be achieved when the random forest model was built based on the features R1, I3, R1G1, R1B1, H1S1, H3I3, L1b1, L3a3, R1G1B1, H1S1I1, or L3a3b3. Among these random forest model models, the distinction performance of the model built based on the feature R1B1 was optimal. For both the SVM model and the random forest model based on the feature R1B1, the distinction accuracies of the training set and testing set were more than 90%, indicating that the feature R1B1 can be utilized to well distinguish between ring rot and anthracnose on apple fruits. In practical applications, the disease image database as a training set may be very large. Generally, the random forest model runs faster than the SVM model, and it is easy to realize large-scale parallel computing and rapid analysis of massive data by using the random forest method. Considering these factors, it is suggested that using R1B1 features to build a random forest model can be carried out for the distinction of the images of ring rot and anthracnose on apple fruits. The results indicated that it is feasible to distinguish between ring rot and anthracnose with typical symptoms on apple fruits by using the method proposed in this study.

Disease images obtained in natural scenes often have complex backgrounds, which can induce uneven illumination, various noises, blurred lesion edges, and other phenomena in the images. Therefore, under these circumstances, it is difficult to conduct complete lesion image segmentation. Zou et al. (2010) used multi-threshold methods to segment apple images from the black background, and then used a flooding algorithm and a snake algorithm to conduct the detection of apple fruit defects, thus the in-line detection of apple quality was realized. The lesion image segmentation method of the two apple fruit diseases based on circle fitting proposed in this study can realize the automatic segmentation of lesion images without human interaction. Even

TABLE 7 | Distinction results of the random forest models based on the LBP histogram features of the gray images of the three components in the L*a*b* color space of the segmented lesion images.

Feature	The number of decision trees built by the best random forest model	Distinction accuracy of the training set/%	Distinction accuracy of the testing set/%
L1	30	100.00	80.49
a1	60	100.00	75.61
b1	100	100.00	78.05
L1a1	60	100.00	82.93
L1b1	70	100.00	85.37
a1b1	40	100.00	80.49
L1a1b1	60	100.00	82.93
L2	10	97.56	70.73
a2	60	100.00	82.93
b2	80	100.00	78.05
L2a2	40	100.00	80.49
L2b2	30	100.00	80.49
a2b2	50	100.00	80.49
L2a2b2	70	100.00	80.49
L3	60	100.00	73.17
a3	20	100.00	80.49
b3	60	100.00	80.49
L3a3	30	100.00	85.37
L3b3	100	100.00	82.93
a3b3	30	100.00	80.49
L3a3b3	80	100.00	87.80

*L1 represents the LBP histogram feature of the gray image of the L^* component of the lesion image filtered by the MB_1-LBP operator; L2 represents the LBP histogram feature of the gray image of the L * component of the lesion image filtered by the MB_2-LBP operator; L1a1 represents the combination of L1 and a1; the rest features' implication could be deduced by analogy.*

if the image background, such as a natural scene with soil, branches, and leaves, is complex, relatively accurate segmentation results can be obtained. Because the strategy of the lesion image segmentation method used in this study was to determine the apple fruit location firstly and then determine the lesion location, the method required that the apple fruits should occupy the main part of the whole image in order to accurately determine the apple fruit location. In this study, it was assumed that the lesions of ring rot and anthracnose on apple fruits were nearly circular, but in practice, two or more lesions may join together, resulting in a great difference between the lesion shape and a circle, which could make the used lesion segmentation method ineffective. In addition, an apple fruit itself is a 3D (three-dimension) object, and a lesion on its surface is a 3D curved surface. The acquired image of the apple fruit and lesion is the projection of the 3D curved surface on a 2D (two-dimension) plane, and an inappropriate camera shooting angle can seriously affect the near circularity after the projection. Therefore, in the process of image capture, the lesions on the apple fruits should be photographed from the front side as much as possible to ensure the accuracy of lesion image segmentation. Meanwhile, the applications of 3D image acquisition and processing technology to plant disease distinction and identification should be strengthened.

Extracted features from the segmented lesion images are the bases of disease image recognition. However, uneven illumination and illumination changes can affect the extracted image features. In this study, the MB-LBP operators were used to extract the texture features in different scales and to reduce the influence of image noise caused by the factors including uneven illumination and illumination changes and based on the extracted features, satisfactory distinction results of the two apple diseases were achieved by using both the SVM method and the random forest method. As a local texture descriptor, the LBP operator has been utilized in plant disease image recognition (Leiva-Valenzuela and Aguilera, 2013; Dubey and Jalal, 2014, 2016; Shrivastava et al., 2017; Araujo and Peixoto, 2019). It has been improved to new operators including CLBP (Dubey and Jalal, 2014, 2016), adaptive center-symmetric local binary patterns (Wang et al., 2016), and square symmetric local binary patterns (Shrivastava et al., 2017) for image recognition of plant diseases. In further studies, more improved LBP operators can be used to explore better feature extraction methods for disease image recognition.

Deep learning has been applied to plant disease image recognition (Tan et al., 2015; DeChant et al., 2017; Lu et al., 2017; Liu et al., 2018; Bansal et al., 2021; Caldeira et al., 2021; Chen et al., 2021; Trivedi et al., 2021; Narmadha et al., 2022). It can reduce image preprocessing operations and achieve satisfactory disease recognition results. Compared with the disease distinction method used in this study, the deep learning method requires a large number of artificially labeled image samples to achieve satisfactory results; otherwise, the trained deep learning model is easy to result in over-fitting and poor generalization ability. In addition, the interpretability of a deep learning model is poor, and it is difficult to determine the reason for a recognition error made by using the model. In this study, the proposed segmentation methods for the images of apple ring rot and apple anthracnose were designed based on visual cognition experiences of human beings, and the purpose of each step involved in the segmentation algorithm is interpretable. When the number of the obtained image samples for training is not enough to meet the requirement of deep learning, the developed method in this study can provide a feasible solution. In further studies, it is expected to explore a general end-to-end apple disease recognition solution based on deep learning and a more comprehensive apple disease image database.

In recent years, image processing and recognition techniques have been developed and applied in many fields such as automatic apple picking (Zhang et al., 2016; Tao and Zhou, 2017; Kang et al., 2020), non-destructive detection of apple fruit quality (Zou et al., 2010; Zhang et al., 2014; Li Y. F. et al., 2021), automatic apples grading (Huang and Fei, 2017; Bhargava and Bansal, 2021), and apple yield estimation (Qian et al., 2013; Li Z. J. et al., 2021). These techniques can be used as references to carry out automatic identification and diagnosis of apple diseases. Computer vision technology can be made full use of to improve the ability of image acquisition and processing and to realize online detection and recognition of apple diseases. Furthermore, the distinction and identification of apple diseases could be implemented by comprehensive utilization of various detection methods. Hyperspectral imaging technology has been used to detect bruises (Xing et al., 2005; Ferrari et al., 2015; Tan et al., 2018; Zhang and Li, 2018) and insect damage (Tian et al., 2015; Rady et al., 2017) on apple fruits. The advantages of hyperspectral imaging technology with the characteristic of combining images with spectra can be used to detect apple diseases (Jarolmasjed et al., 2018; Shuaibu et al., 2018; Solovchenko et al., 2021).

The symptoms of plant diseases may be different at the different growth stages. The images with the typical symptoms of the two apple fruits diseases were used in this study. Further studies on the image distinction and recognition of the two apple fruit diseases with atypical symptoms are needed. Moreover, besides the two apple fruit diseases (ring rot and anthracnose), there are other apple fruit diseases such as scab, fruit moldy core rot caused by *Trichothecium roseum*, and Phytophthora rot caused by *Phytophthora cactorum* (Li B. H. et al., 2013; Hu et al., 2016). The research on image recognition technology for various apple fruit diseases and the different development stages of the diseases should be strengthened. Specially, the research on the early monitoring and early diagnosis of apple fruit diseases based on image processing technology should be carried out for effective and early control of the diseases. An image database of various apple fruit diseases should be established, and the image recognition methods and recognition systems for the diseases should be developed based on image processing technology. With the development and popularity of smartphones, image acquisition becomes more convenient. Apps (mobile applications) related to plant diseases have been developed and applied in practice. Smartphone-based apple disease recognition systems have been reported (Qu et al., 2015), but usually there are a few kinds of apple diseases included, leading to limited applicability. Therefore, the research and development of Apps for image

recognition of various fruit tree diseases should be carried out so that the disease recognition could be conducted in the fields and the problems of the recognition and diagnosis of plant diseases could be solved in time. Furthermore, disease image distinction and identification can be integrated into an apple production management system or an apple post-harvest grading system to better ensure apple production safety and apple product management.

DATA AVAILABILITY STATEMENT

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

AUTHOR CONTRIBUTIONS

HGW contributed to the conception of the study. HGW and QF designed the experiments. QF, SW, HW, ZQ, and HGW performed the experiments. QF and HGW analyzed the data and wrote the draft of the manuscript. All authors contributed to manuscript revision, read, and approved the final version of the manuscript.

FUNDING

This work was supported in part by International Research Exchange Scheme of the Marie Curie Program of the 7th Framework Program (Ref. PIRSES-GA-2013-612659).

REFERENCES

Ahonen, T., Hadid, A., and Pietikäinen, M. (2006). Face description with local binary patterns: application to face recognition. *IEEE Trans. Pattern Anal.* 28, 2037–2041. doi: 10.1109/TPAMI.2006.244

Araujo, J. M. M., and Peixoto, Z. M. A. (2019). A new proposal for automatic identification of multiple soybean diseases. *Comput. Electron. Agric.* 167:105060. doi: 10.1016/j.compag.2019.105060

Bansal, P., Kumar, R., and Kumar, S. (2021). Disease detection in apple leaves using deep convolutional neural network. *Agriculture* 11:617. doi: 10.3390/agriculture11070617

Barbedo, J. G. A. (2014). An automatic method to detect and measure leaf disease symptoms using digital image processing. *Plant Dis.* 98, 1709–1716. doi: 10.1094/PDIS-03-14-0290-RE

Barbedo, J. G. A. (2016). A review on the main challenges in automatic plant disease identification based on visible range images. *Biosyst. Eng.* 144, 52–60. doi: 10.1016/j.biosystemseng.2016.01.017

Bhargava, A., and Bansal, A. (2021). Classification and grading of multiple varieties of apple fruit. *Food Anal. Method* 14, 1359–1368. doi: 10.1007/s12161-021-01970-0

Breiman, L. (2001). Random forests. *Mach. Learn.* 45, 5–32. doi: 10.1023/A:1010933404324

Burges, C. J. C. (1998). A tutorial on support vector machines for pattern recognition. *Data Min. Knowl. Disc.* 2, 121–167. doi: 10.1023/A:1009715923555

Caldeira, R. F., Santiago, W. E., and Teruel, B. (2021). Identification of cotton leaf lesions using deep learning techniques. *Sensors* 21:3169. doi: 10.3390/s21093169

Camargo, A., and Smith, J. S. (2009). Image pattern classification for the identification of disease causing agents in plants. *Comput. Electron. Agric.* 66, 121–125. doi: 10.1016/j.compag.2009.01.003

Chang, C. C., and Lin, C. J. (2011). LIBSVM: a library for support vector machines. *ACM Trans. Intel. Syst. Tech.* 2, 1–27. doi: 10.1145/1961189.1961199

Chen, G. F., Zhao, S., Cao, L. Y., Fu, S. W., and Zhou, J. X. (2021). Corn plant disease recognition based on migration learning and convolutional neural network. *Smart Agric.* 1, 34–44. doi: 10.12133/j.smartag.2019.1.2.201812-SA007

Chen, X. S., Han, M. Y., Su, G. L., Liu, F. Z., Guo, G. N., Jiang, Y. M., et al. (2010). Discussion on today's world apple industry trends and the suggestions on sustainable and efficient development of apple industry in China. *J. Fruit Sci.* 27, 598–604. doi: 10.13925/j.cnki.gsxb.2010.04.038

Chesmore, D., Bernard, T., Inman, A. J., and Bowyer, R. J. (2003). Image analysis for the identification of the quarantine pest *Tilletia indica*. *Bull. OEPP* 33, 495–499. doi: 10.1111/j.1365-2338.2003.00686.x

Cortes, C., and Vapnik, V. (1995). Support-vector networks. *Mach. Learn.* 20, 273–297. doi: 10.1023/A:1022627411411

DeChant, C., Wiesner-Hanks, T., Chen, S. Y., Stewart, E. L., Yosinski, J., Gore, M. A., et al. (2017). Automated identification of northern leaf blight-infected maize plants from field imagery using deep learning. *Phytopathology* 107, 1426–1432. doi: 10.1094/PHYTO-11-16-0417-R

Deng, G. Q., Wang, J. C., Yang, J., Liu, T., Li, D. S., and Sun, C. M. (2021). Identification of Fusarium head blight in wheat ears based on image and improved U-net model. *J. Trit. Crops* 41, 1432–1440. doi: 10.7606/j.issn.1009-1041.2021.11.14

Deng, J. Z., Li, M., Yuan, Z. B., Jin, J., and Huang, H. S. (2012). Feature extraction and classification of *Tilletia* diseases based on image recognition. *Trans. Chin. Soc. Agric. Eng.* 28, 172–176. doi: 10.3969/j.issn.1002-6819.2012.03.030

Dubey, S. R., and Jalal, A. S. (2014). Fusing color and texture cues to identify the fruit diseases using images. *Int. J. Comput. Vis. Image Process.* 4, 52–67. doi: 10.4018/ijcvip.2014040104

Dubey, S. R., and Jalal, A. S. (2016). Apple disease classification using color, texture and shape features from images. *Signal Image Video Process.* 10, 819–826. doi: 10.1007/s11760-015-0821-1

Ferrari, C., Foca, G., Calvini, R., and Ulrici, A. (2015). Fast exploration and classification of large hyperspectral image datasets for early bruise detection on apples. *Chemomet. Intell. Lab.* 146, 108–119. doi: 10.1016/j.chemolab.2015.05.016

Ganthaler, A., Losso, A., and Mayr, S. (2018). Using image analysis for quantitative assessment of needle bladder rust disease of Norway spruce. *Plant Pathol.* 67, 1122–1130. doi: 10.1111/ppa.12842

Gonzalez, R. C., and Woods, R. E. (2011). *Digital Image Processing*, 3rd Edn. Beijing: Publishing House of Electronics Industry.

Halidou, A., You, X. G., and Bogno, B. (2014). Pedestrian detection based on multi-block local binary pattern and biologically inspired feature. *Comput. Inform. Sci.* 7, 125–135. doi: 10.5539/cis.v7n1p125

Hu, Q. Y., Hu, T. L., Wang, Y. N., Wang, S. T., and Cao, K. Q. (2016). Survey on the occurrence and distribution of apple diseases in China. *Plant Prot.* 42, 175–179. doi: 10.3969/j.issn.0529-1542.2016.01032

Huang, C., and Fei, J. Y. (2017). Online apple grading based on decision fusion of image features. *Trans. Chin. Soc. Agric. Eng.* 33, 285–291. doi: 10.11975/j.issn.1002-6819.2017.01.039

Huo, Y. Q., Tang, J. L., Yin, X. Z., and Fang, Y. (2013). Apple disease recognition based on compressive sensing. *Trans. Chin. Soc. Agric. Mach.* 44, 227–232. doi: 10.6041/j.issn.1000-1298.2013.10.036

Jarolmasjed, S., Khot, L. R., and Sankaran, S. (2018). Hyperspectral imaging and spectrometry-derived spectral features for bitter pit detection in storage apples. *Sensors* 18:1561. doi: 10.3390/s18051561

Johannes, A., Picon, A., AlVarez-Gila, A., Echazarra, J., Rodriguez-Vaamonde, S., Navajas, A. D., et al. (2017). Automatic plant disease diagnosis using mobile capture devices, applied on a wheat use case. *Comput. Electron. Agric.* 138, 200–209. doi: 10.1016/j.compag.2017.04.013

Kainat, J., Ullah, S. S., Alharithi, F. S., Alroobaea, R., Hussain, S., and Nazir, S. (2021). Blended features classification of leaf-based cucumber disease using image processing techniques. *Complexity* 2021:9736179. doi: 10.1155/2021/9736179

Kang, H. W., Zhou, H. Y., and Chen, C. (2020). Visual perception and modeling for autonomous apple harvesting. *IEEE Access* 8, 62151–62163. doi: 10.1109/ACCESS.2020.2984556

Kang, M. J., Lee, J. K., and Kang, J. W. (2017). Combining random forest with multi-block local binary pattern feature selection for multiclass head pose estimation. *PLoS One* 12:e0180792. doi: 10.1371/journal.pone.0180792

Karanwal, S. (2021). A comparative study of 14 state of art descriptors for face recognition. *Multimed. Tools Appl.* 80, 12195–12234. doi: 10.1007/s11042-020-09833-2

Leiva-Valenzuela, G. A., and Aguilera, J. M. (2013). Automatic detection of orientation and diseases in blueberries using image analysis to improve their postharvest storage quality. *Food Control* 33, 166–173. doi: 10.1016/j.foodcont.2013.02.025

Li, B. H., Wang, C. X., and Dong, X. L. (2013). Research progress in apple diseases and problems in the disease management in China. *Plant Prot.* 39, 46–54. doi: 10.3969/j.issn.0529-1542.2013.05.007

Li, G. L., Ma, Z. H., and Wang, H. G. (2011). An automatic grading method of severity of single leaf infected with grape downy mildew based on image processing. *J. China Agric. Univ.* 16, 88–93. doi: 10.11841/j.issn.1007-4333.2011.06.014

Li, G. L., Ma, Z. H., and Wang, H. G. (2012). Image recognition of wheat stripe rust and wheat leaf rust based on support vector machine. *J. China Agric. Univ.* 17, 72–79. doi: 10.11841/j.issn.1007-4333.2012.02.013

Li, H. L., and He, D. J. (2013). Study on technology of restore and location of apples under occluded. *J. Agric. Mech. Res.* 35, 20–23. doi: 10.13427/j.cnki.njyi.2013.09.016

Li, X. L., Ma, Z. H., Bienvenido, F., Qin, F., Wang, H. G., and Alvarez-Bermejo, J. A. (2017). Development of automatic counting system for urediospores of wheat stripe rust based on image processing. *Int. J. Agric. Biol. Eng.* 10, 134–143. doi: 10.25165/j.ijabe.20171005.3084

Li, X. L., Ma, Z. H., Sun, Z. Y., and Wang, H. G. (2013). Automatic counting for trapped urediospores of *Puccinia striiformis* f. sp. *tritici* based on image processing. *Trans. Chin. Soc. Agric. Eng.* 29, 199–206. doi: 10.3969/j.issn.1002-6819.2013.02.028

Li, X. L., Ruan, Q. Q., Jin, Y., An, G. Y., and Zhao, R. Z. (2015). Fully automatic 3D facial expression recognition using polytypic multi-block local binary patterns. *Signal Process.* 108, 297–308. doi: 10.1016/j.sigpro.2014.09.033

Li, Y. F., Feng, X. Y., Liu, Y. D., and Han, X. C. (2021). Apple quality identification and classification by image processing based on convolutional neural networks. *Sci. Rep.* 11:16618. doi: 10.1038/s41598-021-96103-2

Li, Z. J., Yang, S. H., Shi, D. S., Liu, X. X., and Zheng, Y. J. (2021). Yield estimation method of apple tree based on improved lightweight YOLOv5. *Smart Agric.* 3, 100–114. doi: 10.12133/j.smartag.2021.3.2.202105-SA005

Liao, S. C., Zhu, X. X., Lei, Z., Zhang, L., and Li, S. Z. (2007). Learning multi-scale block local binary patterns for face recognition. *Lect. Notes Comput. Sci.* 4642, 828–837. doi: 10.1007/978-3-540-74549-5_87

Liu, B., Zhang, Y., He, D. J., and Li, Y. X. (2018). Identification of apple leaf diseases based on deep convolutional neural networks. *Symmetry* 10:11. doi: 10.3390/sym10010011

Lu, J. W., Liong, V. E., and Zhou, J. (2018). Simultaneous local binary feature learning and encoding for homogeneous and heterogeneous face recognition. *IEEE Trans. Pattern Anal.* 40, 1979–1993. doi: 10.1109/TPAMI.2017.2737538

Lu, Y., Yi, S. J., Zeng, N. Y., Liu, Y. R., and Zhang, Y. (2017). Identification of rice diseases using deep convolutional neural networks. *Neurocomputing* 267, 378–384. doi: 10.1016/j.neucom.2017.06.023

Nachtigall, L. G., Araujo, R. M., and Nachtigall, G. R. (2017). Use of images of leaves and fruits of apple trees for automatic identification of symptoms of diseases and nutritional disorders. *Int. J. Monit. Surveill. Tech. Res.* 5, 1–14. doi: 10.4018/IJMSTR.2017040101

Nanni, L., Lumini, A., and Brahnam, S. (2012). Survey on LBP based texture descriptors for image classification. *Expert Syst. Appl.* 39, 3634–3641. doi: 10.1016/j.eswa.2011.09.054

Narmadha, R. P., Sengottaiyan, N., and Kavitha, R. J. (2022). Deep transfer learning based rice plant disease detection model. *Intell. Autom. Soft Comput.* 31, 1257–1271. doi: 10.32604/iasc.2022.020679

Oberti, R., Marchi, M., Tirelli, P., Calcante, A., Iriti, M., and Borghese, A. N. (2014). Automatic detection of powdery mildew on grapevine leaves by image analysis: optimal view-angle range to increase the sensitivity. *Comput. Electron. Agric.* 104, 1–8. doi: 10.1016/j.compag.2014.03.001

Ojala, T., Pietikäinen, M., and Harwood, D. (1996). A comparative study of texture measures with classification based on feature distributions. *Pattern Recogn.* 29, 51–59. doi: 10.1016/0031-3203(95)00067-4

Omrani, E., Khoshnevisan, B., Shamshirband, S., Saboohi, H., Anuar, N. B., and Nasir, M. H. N. M. (2014). Potential of radial basis function-based support vector regression for apple disease detection. *Measurement* 55, 512–519. doi: 10.1016/j.measurement.2014.05.033

Ortega-Sánchez, N., Rodríguez-Esparza, E., Oliva, D., Pérez-Cisneros, M., Mohamed, A. W., Dhiman, G., et al. (2022). Identification of apple diseases in digital images by using the Gaining-sharing knowledge-based algorithm for multilevel thresholding. *Soft Comput.* 26, 2587–2623. doi: 10.1007/s00500-021-06418-5

Panda, R., Puhan, N. B., Rao, A., Padhy, D., and Panda, G. (2018). Automated retinal nerve fiber layer defect detection using fundus imaging in glaucoma. *Comput. Med. Imag. Grap.* 66, 56–65. doi: 10.1016/j.compmedimag.2018.02.006

Phadikar, S., Sil, J., and Das, A. K. (2013). Rice diseases classification using feature selection and rule generation techniques. *Comput. Electron. Agric.* 90, 76–85. doi: 10.1016/j.compag.2012.11.001

Pires, R. D. L., Gonçalves, D. N., Oruê, J. P. M., Kanashiro, W. E. S., Rodrigues, J. F., Machado, B. B., et al. (2016). Local descriptors for soybean disease recognition. *Comput. Electron. Agric.* 125, 48–55. doi: 10.1016/j.compag.2016.04.032

Powers, D. M. W. (2011). Evaluation: from precision, recall and *F*-measure to ROC, informedness, markedness & correlation. *J. Mach. Learn. Tech.* 2, 37–63.

Pydipati, R., Burks, T. F., and Lee, W. S. (2006). Identification of citrus disease using color texture features and discriminant analysis. *Comput. Electron. Agric.* 52, 49–59. doi: 10.1016/j.compag.2006.01.004

Qian, J. P., Li, M., Yang, X. T., Wu, B. G., Zhang, Y., and Wang, Y. A. (2013). Yield estimation model of single tree of Fuji apples based on bilateral image identification. *Trans. Chin. Soc. Agric. Eng.* 29, 132–138. doi: 10.3969/j.issn.1002-6819.2013.11.017

Qin, F., Liu, D. X., Sun, B. D., Ruan, L., Ma, Z. H., and Wang, H. G. (2016). Identification of alfalfa leaf diseases using image recognition technology. *PLoS One* 11:e0168274. doi: 10.1371/journal.pone.0168274

Qu, Y., Tao, B., Wang, Z. J., and Wang, S. T. (2015). Design of apple leaf disease recognition system based on Android. *J. Agric. Univ. Hebei* 38, 102–106. doi: 10.13320/j.cnki.jauh.2015.0144

Rady, A., Ekramirad, N., Adedeji, A. A., Li, M., and Alimardani, R. (2017). Hyperspectral imaging for detection of codling moth infestation in GoldRush apples. *Postharvest Biol. Technol.* 129, 37–44. doi: 10.1016/j.postharvbio.2017.03.007

Sankaran, S., Maja, J. M., Buchanon, S., and Ehsani, R. (2013). Huanglongbing (citrus greening) detection using visible, near infrared and thermal imaging techniques. *Sensors* 13, 2117–2130. doi: 10.3390/s130202117

Sankaran, S., Mishra, A., Ehsani, R., and Davis, C. (2010). A review of advanced techniques for detecting plant diseases. *Comput. Electron. Agric.* 72, 1–13. doi: 10.1016/j.compag.2010.02.007

Shrivastava, S., Singh, S. K., and Hooda, D. S. (2015). Color sensing and image processing-based automatic soybean plant foliar disease severity detection and estimation. *Multimed. Tools Appl.* 74, 11467–11484. doi: 10.1007/s11042-014-2239-0

Shrivastava, S., Singh, S. K., and Hooda, D. S. (2017). Soybean plant foliar disease detection using image retrieval approaches. *Multimed. Tools Appl.* 76, 26647–26674. doi: 10.1007/s11042-016-4191-7

Shuaibu, M., Lee, W. S., Schueller, J., Gader, P., Hong, Y. K., and Kim, S. (2018). Unsupervised hyperspectral band selection for apple Marssonina blotch detection. *Comput. Electron. Agric.* 148, 45–53. doi: 10.1016/j.compag.2017.09.038

Solovchenko, A., Dorokhov, A., Shurygin, B., Nikolenko, A., Velichko, V., Smirnov, I., et al. (2021). Linking tissue damage to hyperspectral reflectance for non-invasive monitoring of apple fruit in orchards. *Plants* 10:310. doi: 10.3390/plants10020310

Tan, W. X., Zhao, C. J., Wu, H. R., and Gao, R. H. (2015). A deep learning network for recognizing fruit pathologic images based on flexible momentum. *Trans. Chin. Soc. Agric. Mach.* 46, 20–25. doi: 10.6041/j.issn.1000-1298.2015.01.004

Tan, W. Y., Sun, L. J., Yang, F., Che, W. K., Ye, D. D., Zhang, D., et al. (2018). Study on bruising degree classification of apples using hyperspectral imaging and GS-SVM. *Optik* 154, 581–592. doi: 10.1016/j.ijleo.2017.10.090

Tao, Y. T., and Zhou, J. (2017). Automatic apple recognition based on the fusion of color and 3D feature for robotic fruit picking. *Comput. Electron. Agric.* 142, 388–396. doi: 10.1016/j.compag.2017.09.019

Tian, Y. W., Cheng, Y., Wang, X. Q., and Liu, S. J. (2015). Recognition method of insect damage and stem/calyx on apple based on hyperspectral imaging. *Trans. Chin. Soc. Agric. Eng.* 31, 325–331. doi: 10.3969/j.issn.1002-6819.2015.04.046

Tian, Y. W., Li, T. L., Li, C. H., Piao, Z. L., Sun, G. K., and Wang, B. (2007). Method for recognition of grape disease based on support vector machine. *Trans. Chin. Soc. Agric. Eng.* 23, 175–180. doi: 10.3321/j.issn:1002-6819.2007.06.034

Trivedi, N. K., Gautam, V., Anand, A., Aljahdali, H. M., Villar, S. G., Anand, D., et al. (2021). Early detection and classification of tomato leaf disease using high-performance deep neural network. *Sensors* 21:7987. doi: 10.3390/s21237987

Vakilian, K. A., and Massah, J. (2013). An artificial neural network approach to identify fungal diseases of cucumber (*Cucumis sativus* L.) plants using digital image processing. *Arch. Phytopathol. Plant Prot.* 46, 1580–1588. doi: 10.1080/03235408.2013.772321

Vieira, R. A., Mesquini, R. M., Silva, C. N., Hata, F. T., Tessmann, D. J., and Scapim, C. A. (2014). A new diagrammatic scale for the assessment of northern corn leaf blight. *Crop Prot.* 56, 55–57. doi: 10.1016/j.cropro.2011.04.018

Vishnoi, V. K., Kumar, K., and Kumar, B. (2021). Plant disease detection using computational intelligence and image processing. *J. Plant Dis. Protect.* 128, 19–53. doi: 10.1007/s41348-020-00368-0

Wang, X. F., Zhang, S. W., and Kong, W. W. (2016). Adaptive center-symmetric local binary patterns for crop disease recognition. *Guangdong Agric. Sci.* 43, 140–145. doi: 10.16768/j.issn.1004-874X.2016.09.021

Wang, X. P., Huang, T., Tan, W. X., Wu, H. R., and Sun, C. (2015). Apple lesion image recognition method based on convolutional network. *Comput. Eng.* 41, 293–298. doi: 10.3969/j.issn.1000-3428.2015.12.055

Wang, Y., Chen, D. R., Shen, M. L., and Wu, G. (2008). Watershed segmentation based on morphological gradient reconstruction and marker extraction. *J. Image Graph.* 13, 2176–2180. doi: 10.11834/jig.20081119

Wang, Y. F., Mao, H. P., Zhang, X. D., Liu, Y., and Du, X. X. (2021). A rapid detection method for tomato gray mold spores in greenhouse based on microfluidic chip enrichment and lens-less diffraction image processing. *Foods* 10:3011. doi: 10.3390/foods10123011

Xing, J., Bravo, C., Jancsok, P. T., Ramon, H., and Baerdemaeker, J. D. (2005). Detecting bruises on 'Golden Delicious' apples using hyperspectral imaging with multiple wavebands. *Biosyst. Eng.* 90, 27–36. doi: 10.1016/j.biosystemseng.2004.08.002

Yamamoto, K., Togami, T., and Yamaguchi, N. (2017). Super-resolution of plant disease images for the acceleration of image-based phenotyping and vigor diagnosis in agriculture. *Sensors* 17:2557. doi: 10.3390/s17112557

Yang, B., and Chen, S. C. (2013). A comparative study on local binary pattern (LBP) based face recognition: LBP histogram versus LBP image. *Neurocomputing* 120, 365–379. doi: 10.1016/j.neucom.2012.10.032

Yin, X. Z., He, D. J., and Huo, Y. Q. (2012). The intelligent identification methods of apple fruit's disease in nature outdoors based on low-resolution image. *J. Agric. Mech. Res.* 34, 29–32. doi: 10.13427/j.cnki.njyi.2012.10.016

Zhang, B., Su, J. X., Zhang, W. W., Deng, M. H., and Wang, X. Z. (2016). Design for apple-picking robot of intelligent recognition based on laser vision. *J. Agric. Mech. Res.* 38, 60–64. doi: 10.13427/j.cnki.njyi.2016.07.012

Zhang, B. H., Huang, W. Q., Li, J. B., Zhao, C. J., Liu, C. L., and Huang, D. F. (2014). On-line identification of defect on apples using lightness correction and AdaBoost methods. *Trans. Chin. Soc. Agric. Mach.* 45, 221–226. doi: 10.6041/j.issn.1000-1298.2014.06.034

Zhang, C. L., Zhang, S. W., Yang, J. C., Shi, Y. C., and Chen, J. (2017). Apple leaf disease identification using genetic algorithm and correlation based feature selection method. *Int. J. Agr. Biol. Eng.* 10, 74–83. doi: 10.3965/j.ijabe.20171002.2166

Zhang, L., Chu, R. F., Xiang, S. M., Liao, S. C., and Li, S. Z. (2007). Face detection based on multi-block LBP representation. *Lect. Notes Comput. Sci.* 4642, 11–18. doi: 10.1007/978-3-540-74549-5_2

Zhang, M., and Li, G. H. (2018). Visual detection of apple bruises using AdaBoost algorithm and hyperspectral imaging. *Int. J. Food Prop.* 21, 1598–1607. doi: 10.1080/10942912.2018.1503299

Zhang, Q., and Wang, Z. L. (2009). *Mastering MATLAB Image Processing*. Beijing: Publishing House of Electronics Industry.

Zhang, S. W., Wu, X. W., You, Z. H., and Zhang, L. Q. (2017). Leaf image based cucumber disease recognition using sparse representation classification. *Comput. Electron. Agric.* 134, 135–141. doi: 10.1016/j.compag.2017.01.014

Zhang, Z., Ni, H. X., Yuan, C. M., and Yang, L. H. (2013). *Mastering MATLAB Digital Image Processing and Recognition*. Beijing: Posts & Telecom Press.

Zhu, J. H., Wu, A., Wang, X. S., and Zhang, H. (2020). Identification of grape diseases using image analysis and BP neural networks. *Multimed. Tools Appl.* 79, 14539–14551. doi: 10.1007/s11042-018-7092-0

Zou, X. B., Zhao, J. W., Li, Y. X., and Holmes, M. (2010). In-line detection of apple defects using three color cameras system. *Comput. Electron. Agric.* 70, 129–134. doi: 10.1016/j.compag.2009.09.014

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

frontiers | Frontiers in Plant Science

METHODS
published: 09 June 2022
doi: 10.3389/fpls.2022.888978

Quality Control System of Red Jujube by Hybrid Model: Development of an Efficient Framework

*Rongzhi Jing and Ping Li**

School of Electronic and Information Engineering, Sias University, Xinzheng, China

Food traceability is very important for the quality and safety of agricultural products, which is related to the people's livelihood and national economy and has drawn great attention from governments and scientists around the world. The existing studies have not yet overcome the crisis characteristics comprehensively and systematically. A traceability system of red jujube is constructed by a hybrid mode of blockchain and the Internet of Things (IoTs). The system integrates the blockchain and the IoT technologies with characteristics of tamper-proof, decentralization, and distributed storage and solves the problem of date quality traceability by designing the technical process and architecture of date quality traceability and the big data of red jujube, jujube plantation, processing enterprise, commercial enterprises, and market administration. The whole process from planting to processing and sales of red jujube are recorded in the block to ensure the realization of quality traceability of red dates in the process. Through the whole process of big data processing, the key information collected in each process is stored in the database to ensure the realization of quality traceability of red dates in the framework. The system can help to minimize the production and distribution of unsafe or poor-quality products, thereby minimizing the potential for bad publicity, liability, and recalls.

Keywords: red jujube quality and safety, traceability system, blockchain, internet of things (IoTs), quality control system

OPEN ACCESS

Edited by:
Chuanlei Zhang,
Tianjin University of Science and Technology, China

Reviewed by:
Catello Di Martino,
University of Molise, Italy
Saeed Hamood Alsamhi,
Ibb University, Yemen

***Correspondence:**
Ping Li
19360666@qq.com

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

Received: *03 March 2022*
Accepted: *10 May 2022*
Published: *09 June 2022*

Citation:
Jing R and Li P (2022) Quality Control System of Red Jujube by Hybrid Model: Development of an Efficient Framework. Front. Plant Sci. 13:888978. doi: 10.3389/fpls.2022.888978

INTRODUCTION

The quality and safety of agricultural products are related to the people's livelihood and national economy. The frequent occurrence of agricultural product quality and safety crises has drawn great attention from countries around the world. Quality control in agricultural production and throughout the supply chain is one of the most challenging issues in the world today, especially when it comes to typical food products. The quality and safety of agricultural products are an important study direction at present, which is a multidisciplinary study field. Its purpose is to explore the technology, measure and supervision guarantee system to ensure the hygiene and food safety of agricultural products, reduce the potential disease, and prevent the risk of food poisoning in the process of agricultural products processing, storage, and sales. In recent years, many agricultural quality control systems have been presented and various companies are trying to experiment with joint use of radiofrequency identification (RFID), Internet of Things (IoTs), and blockchain technologies to solve problems in scenarios where numerous untrusted actors get involved (Galvez et al., 2018; Corallo et al., 2020).

Good traceability systems help to minimize the production and distribution of unsafe or poor quality products and minimize the potential for bad publicity, liability, and recalls. The current

food labeling system cannot guarantee that the food is authentic, good quality, and safe. Zhao et al. (2019) presented comprehensive information about traceability with regard to safety and quality in the food supply chain and applied as a tool to assist in the assurance of food safety and quality, as well as to achieve consumer confidence.

Blockchain technology is a new digital technological approach underpinned by the Industry 4.0. It is used to ensure data integrity and prevent tampering and single point of failure through offering fault tolerance, immutability, trust, transparency, and full traceability of the stored transaction records to all the agri-food value chain partners. Zhao et al. (2019) used systematic literature network analysis to review the state-of-the-art blockchain technology, including its recent advances, main applications in agrifood value chain, and challenges from a holistic perspective. They identified six challenges that have been included storage capacity and scalability, privacy leakage, high cost and regulation problem, throughput and latency issue, and lack of skills. Yadav et al. (2020) identified seven enablers by grouping thirty subenablers, developed the Internet of Things (IoTs)-based efficient and supportive coordinating system for enhancing the coordinating mechanism in Agriculture Supply Chain Management (ASCM) under natural outbreaks and discussed a case study of the sugar mill industry. Zu et al. (2020) analyzed and researched in detail the typical crisis events of agricultural product quality and safety in China from 2004 to 2018. They extracted 13 abstract features to form a set of agricultural product quality and safety crisis features. Their studies are helpful to enrich the study results of quality and safety management of agricultural products in China and take effective measures to reduce the harm. Rakitskii et al. (2020) quantified 42 active ingredients of pesticides (and their metabolites) in samples of rice grain, dragon fruit (pitahaya), avocado, mango, and banana (fresh and dried) and used the sample preparation procedure with QuEChERS technology for the multiresidues determination of the 40 compounds. The results show the safety of certain types of food products imported from Vietnam by the content of residual quantities of pesticides. Demestichas et al. (2020) overviewed the application of blockchain technologies for enabling traceability in the agri-food domain and presented definitions, levels of adoption, tools, and advantages of traceability, accompanied with a brief overview of the functionality and advantages of blockchain technology. They conducted an extensive literature review on the integration of blockchain into traceability systems and discussed the relevant existing commercial applications, highlighting the relevant challenges and future prospects of the application of blockchain technologies in the agri-food supply chain. Feng et al. (2020) reviewed the blockchain technology characteristics and functionalities, provided valuable information for researchers and practitioners on the use of blockchain-based food traceability management and proposed an architecture design framework and suitability application analysis flowchart of blockchain-based food traceability systems. By combining smart contracts, Interplanetary File System, and the Internet of Things, Cocco and Mannaro (2021) tried to address these issues and presented a proposal of an implementation model for the supply chain management of a typical Italian Carasau bread. The method can guarantee and certify a transparent, secure, and auditable traceability in such a way that each actor of the supply chain can verify the quality of the product. Ray et al. (2021) looked into manufacturer–retailer collaboration in the UK food supply chain and aimed to develop a preliminary conceptual framework by identifying the key factors that influence long-term performance and accuracy of collaborative forecasting and developed a new model for Synchronized Information Forecast Collaboration (SIFC). Zhao and Ning (2017) constructed an agricultural traceability system of agricultural products through the IoT technologies of information, security, and cloud computing. This system can provide effective basis for agricultural trade, logistics, and safe consumption. The customers can effectively understand the detailed process and risk status of every phase of agricultural products through this system. Almalki et al. (2021) presented a low-cost platform for comprehensive environmental parameter monitoring using the IoT to help farmers, government, or manufacturers to predict environmental data over the geographically large farm field, which can enhance crop productivity and farm management in a cost-effective and timely manner. Yang and Sun (2020) introduced blockchain technology and proposed a blockchain-based data management system to afford efficient data extraction, management, and access control for heterogeneous forms of data across the agricultural supply chain. Zhang et al. (2020) analyzed the operation mechanism and development path of agricultural product supply chain by blockchain technology and reconstructed the agricultural product supply chain based on the advantages of blockchain technology. Blockchain makes data public for all the drones and enables drones to log information concerning world states, time, location, resources, delivery data, and drone relation to all the neighbor drones. Alsamhi et al. (2021) introduced decentralized independent multidrones to accomplish the task collaboratively and discussed end-to-end delivery application of combination of blockchain and multidrone in combating coronavirus disease 2019 (COVID-19) and beyond future pandemics. Edge computing has prospects in agricultural applications, such as safety traceability of agricultural products, pest identification, unmanned agricultural machinery, and intelligent management. It is possible to apply federated learning to beyond 5G by development of edge computing makes. Alsamhi et al. (2022) proposed a blockchain empowered federated learning framework, presented its potential application scenarios in beyond 5G, and designed a deep reinforcement learning-based algorithm to find an optimal solution to the problem. The results showed that the proposed scheme is effective. Zhang et al. (2021) primarily reviewed the application of edge computing in the agricultural IoTs and investigated the combination of edge computing and artificial intelligence, blockchain, and virtual/augmented reality technology. Tharatipyakul and Pongnumkul (2021) gathered 25 review articles on blockchain or agri-food supply chain and 39 study articles that presented screenshots of user interfaces of related applications, reviewed 7 review articles that focused on the blockchain-based agri-food supply chain to understand the benefits and challenges in the blockchain applications, aimed

to address this gap by reviewing existing works from user interface perspectives, and analyzed 14 blockchain-based agri-food traceability applications and 10 non-blockchain-based agri-food traceability applications. Finally, they discussed the study gaps and future study directions related to user interface design, which should be addressed to ease future blockchain adoption.

Agricultural products traceability management integrated system can produce the data uploaded to the agricultural products of the whole process of cultivation of traceability management integrated system, the digital information, so supervision department can do random inspection of agricultural products, through the data to verify reliability, record results, give consumers a greater sense of trust, enhance consumer confidence. The problems of traceability agricultural products system are summarized as: (1) Enterprises developing traceability agricultural products system are not the same, so traceability information cannot be shared, many system software are not compatible, and the purpose of query cannot be achieved in different systems. Some query terminals can only be queried in supermarkets, which make it inconvenient to use; (2) At present, the traceability system is not unified in many aspects, such as identification code, storage information, and network query system, and it also faces different kinds of food; and (3) The accuracy of traceability is not high, some of them can only be traced back to the enterprise, and the specific process of planting, processing, or transportation cannot be traced.

Xinzheng red jujube (XRJ) is a specialty of Xinzheng city, Zhengzhou City, Henan Province, China. It is a good fruit nourishing blood and spleen beauty, has high medicinal value, and its leaves, flowers, fruits, skins, roots, and thorns can be used as medicine. However, the development of XRJ industry is also facing many problems, such as the quality and safety system is not perfect, especially the pesticide and fertilizer residues are too high, which greatly restrict the development pace of XRJ industry. The quality and safety of XRJ is not only a major livelihood issue related to people's health, but also a major obstacle to the international XRJ trade. To guarantee safety in food, an efficient tracking and tracing system is required. RFID devices allow recording all the useful information for traceability directly on the commodity. By analyzing the problems existing in the quality and safety of XRJ and discussing the quality and safety problems of XRJ based on the IoT, this article tries to construct a quality control system of XRJ based on the IoT, improve the quality and safety control level of XRJ, and then realize the intelligent and scientific management of XRJ production, which have important practical significance for promoting the development of modern XRJ. This proposed system aims to reduce the chemical pesticides, fertilizers and antibiotics usage, stimulating sustainable food consumption, and promoting affordable healthy food for all. Consumers are encouraged to choose healthy and sustainable diets and reduce food waste. Farmers and producers are encouraged to provide more details about food origin, nutritional value, and environmental footprint. The contributions of this article are as follows.

(1) A hybrid mode of blockchain and the IoT is used to construct a quality control system of XRJ.
(2) Data processing of XRJ traceability system is introduced in detail.
(3) The XRJ traceability system is verified in the Henan Xinzheng Xinxing Jujube Industry Corporation Ltd.

The rest of the article is organized as follows. In Section 2, the related work is introduced. Section 3 focuses on XRJ quality traceability system in detail. The experiments and analysis are conducted in Section 4. Section 5 concludes the study work and point out the future work.

RELATED WORK

Agricultural Internet of Things

The Internet of Things (IoTs) is an extended network based on the internet through a variety of sensors, RFID technology, global positioning system (GPS), infrared sensors, laser scanner, and other equipment and technology, no need to monitor real-time acquisition, connected, interactive object or process, collect the sound, light, heat, electricity, mechanics, chemistry, biology, location, and other needed information, through all the kinds of possible internet access, to realize the ubiquitous connection between objects and objects, objects and people, and realize the intelligent perception, recognition, and management of objects and processes. The Internet of Things is an information carrier based on the internet, traditional telecommunications networks, etc. It enables all the ordinary physical objects that can be independently addressed to form an interconnected network (Mukkamala et al., 2018).

At present, XRJ production mainly depends on manual experience management and lack of scientific and systematic guidance. Information acquisition is one of the most important key technologies to realize the high level of facility production of XRJ and to optimize the facility biological environment control. The agricultural IoT plays a crucial role in revolutionizing agricultural production. From the practical experience of developed countries, the IoT has many applications in the field of agriculture and is developing in the direction of reliability, energy saving, environmental adaptability, low cost, information, and intelligence. If it is a jujube production base of thousands of acres, if the IoT technology is applied, manual control only needs to click the mouse tiny action, only a few seconds before and after, can completely replace the tedious manual operation such as watering and fertilization. From the point of view of different stages of XRJ production, whether from the cultivation stage of planting and harvest stage, the IoT technology can be used to improve its work efficiency and fine management.

- At the plant preparation stage, a bunch of sensors can be arranged in the greenhouse to analyze real-time soil information and choose the right date varieties.
- At the planting and cultivation stage of jujube, the IoT technology can be used to collect temperature and humidity information for efficient management, so as to cope with environmental changes and to ensure plant seedlings grow in the best environment.

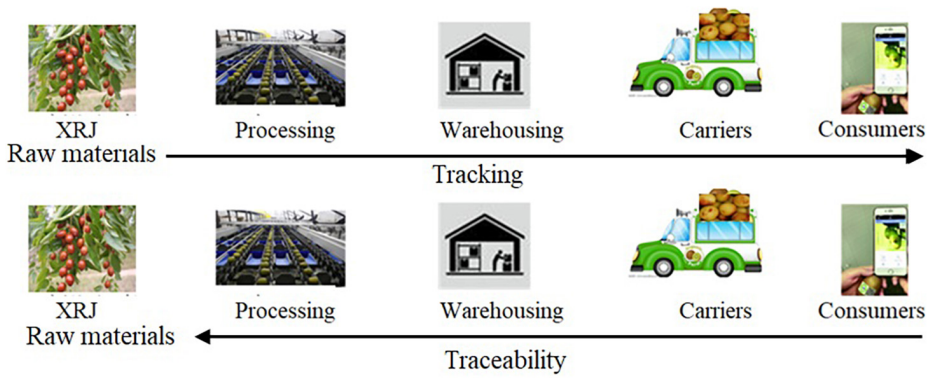


FIGURE 1 | Tracking and traceability system of Xinzheng red jujube (XRJ).

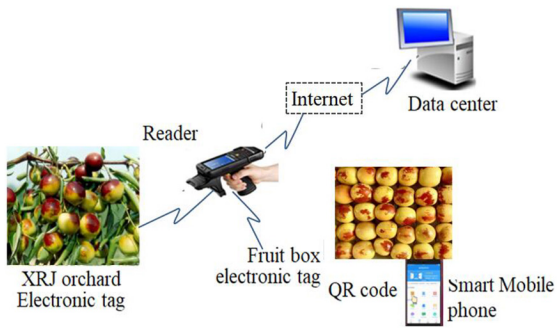


FIGURE 2 | XRJ tracing query based on radiofrequency identification (RFID) + Quick Response (QR) code.

- At the growth stage of jujube, the IoT can be used to monitor the environmental information, nutrient information, and crop diseases and insect pests in real time. Using relevant sensors are accurate and real-time access to soil moisture, environmental temperature and humidity, illumination, through real-time data monitoring and special varieties of expert experience, combining with control system regulating crop growth environment, improve crop nutrition state, crop pest and disease outbreaks in time period, to maintain the best crop growth conditions. It plays a very important role in the growth and management of jujube.
- At the harvest stage of jujube, the IoT technology is used to transfer information such as appearance, fruit diameter, and nutritional composition of jujube to the data center. According to the data of planting environment, weather condition, and fruit tree status, the planting scheme is optimized for next year through big data technology analysis.

Through the IoT technology, the production and operation of XRJ can be rapidly transformed from extensive and empirical management to fine and scientific management, so as to improve the yield and quality of XRJ, reduce the production cost of XRJ, and protect the agricultural environment. **Figure 1** shows the flowchart of tracking and traceability system, where tracking is the ability to follow the path of a particular unit or batch of XRJ from upstream to downstream of the supply chain and traceability is the ability to identify the source of a specific unit or batch of XRJ products from the downstream to the upstream of the supply chain, i.e., the ability to trace the planting, production, and processing of a certain XRJ product through the method of traceability code.

Data acquisition in **Figure 1** is implemented by RFID + Quick Response (QR) code. **Figure 2** shows the XRJ tracing query flowchart based on RFID + QR code.

Blockchain

Blockchain is a chain of blocks, each containing a certain amount of information, connected in the chronological order of their creation (Konstantinidis et al., 2018). The chain is stored on all the servers and as long as one server in the system works, the whole blockchain is safe. These servers, called nodes in a blockchain system, provide storage space and computing power for the entire blockchain system. Tampering with information in a blockchain is extremely difficult because it requires the consent of more than half of the nodes and modification of all the nodes, which are usually in the hands of different parties. Compared with traditional networks, blockchain has two core features: data are hard to tamper with and it is decentralized. Based on these two characteristics, the information recorded by blockchain is more authentic and reliable, which can help to solve the problem of distrust among people. In the narrow sense, blockchain is a distributed ledger that cannot be tampered with or forged by means of cryptography. It is a chain data structure that combines data blocks sequentially according to time sequence. Generalized blockchain technology is the use of blockchain data validation and data storage structure and update the data generated by distributed node consensus algorithm, using the way to ensure the safety of data transmission and access of cryptography, the use of automated script code intelligent contracts, programming and operation data of new distributed infrastructure, and computing paradigm.

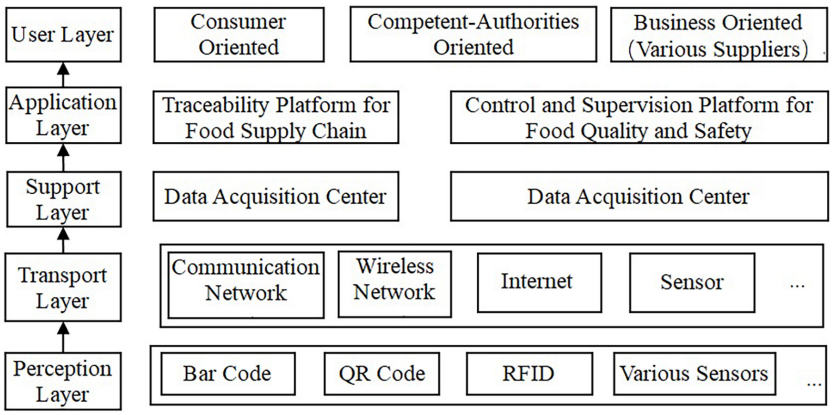


FIGURE 3 | The architecture of XRJ traceability system.

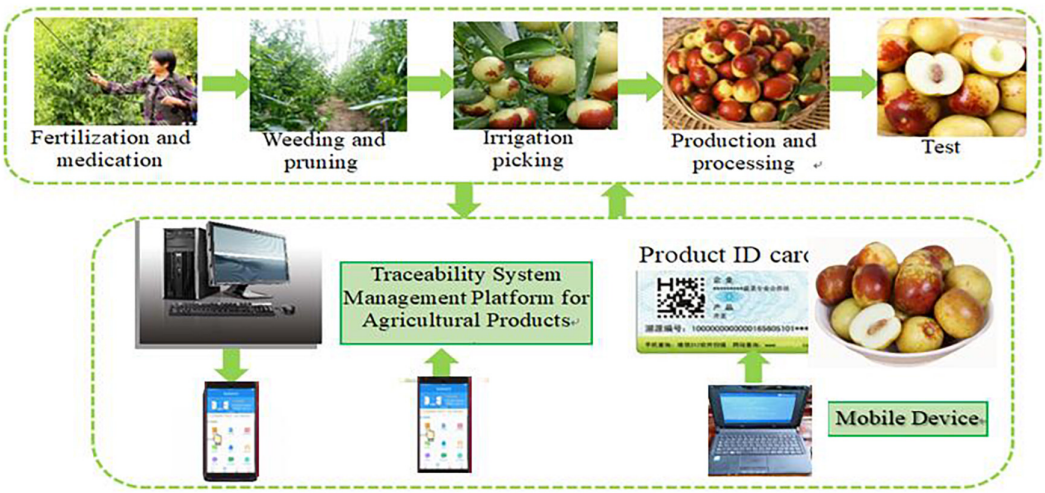


FIGURE 4 | The flowchart of XRJ traceability system.

Blockchains allow people to have a distributed peer-to-peer network where non-trusting members can interact with each other without a trusted intermediary in a verifiable manner. Combining blockchain with the IoT, namely, blockchain-IoT is powerful and can cause significant transformations across several industries, paving the way for new business models and novel, distributed applications (Christidis and Devetsikiotis, 2016). It facilitates the sharing of services and resources leading to the creation of a marketplace of services between devices and allows people to automate in a cryptographically verifiable manner several existing, time-consuming workflows.

XINZHENG RED JUJUBE QUALITY TRACEABILITY SYSTEM

From the actual needs of the whole quality and safety management of XRJ, based on the quality and safety standards of XRJ, technical standards of XRJ, safety law, and other regulations and standards, the whole process risk management and traceability system framework of pollution-free XRJ quality and safety are designed. The wireless network remote environment monitoring technology is used to collect and analyze the data of XRJ in the whole circulation links such as XRJ, processing, storage, transportation, and sales, and generate reports with the results, so that managers can understand the situation in each link and solve related problems in time. All the monitoring report and pictures are numbered, using the information classification and coding standards, designing the pollution-free XRJ quality security coding, generating the QR code identification of XRJ, so that the consumers by smartphone scanning, back to the whole process of XRJ circulation. The whole process of pollution-free XRJ quality and safety can be monitored and traced.

The XRJ quality traceability system based on hybrid mode is mainly divided into hardware and software system. The hardware module mainly realizes the real-time data collection of crops from planting to processing, circulation and sales of XRJ, and writes the data into the database. The software system is divided

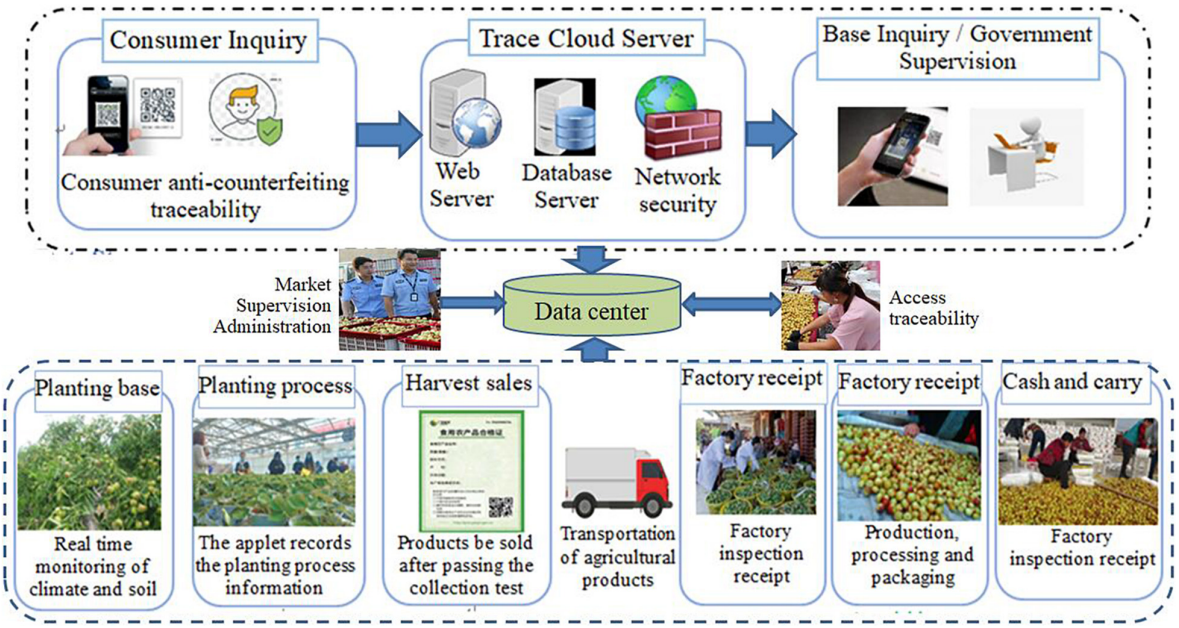


FIGURE 5 | Data processing of XRJ traceability system.

FIGURE 6 | Implementation processing of XRJ traceability system.

into Web Service and Android client. Since the Android client cannot interact directly with the database, this document uses Web Service as a bridge. The Web Service retrieves data from the database and the Android client retrieves data from the Web Service through Simple Object Access Protocol (SOAP).

(1) Hardware environment design: The hardware part of the system mainly includes power supply module, sensor module, ZigBee module, and gateway module. The power supply module supplies power to other modules. The sensor module collects environmental data in real time and then transmits the data to the gateway module through ZigBee module. Finally, the gateway module uploads the data to the database through a specific serial port communication program.

(2) Software design: The gateway module of the hardware part writes the real-time data collected by the sensor into the database through serial communication. The Web Service obtains data from the database and the Android client obtains data from the Web Service through SOAP protocol. The Android client interacts with the database through Web Service, so as to realize the data query function.

The architecture and flowchart of traceability system of XRJ based on hybrid mode of blockchain and the IoT are shown in **Figures 3**, **4**, respectively.**Figure 5** shows the data processing process of the system based on hybrid mode of blockchain and the IoT, including network data process and business process. **Figure 6** shows the implementation processing. Data can be stored in two ways: database and blockchain. The database is maintained by the Production and market supervision administration in the traceability system, while the blockchain is jointly maintained by jujube farmers, jujube gardens, industrial

TABLE 1 | System development environment configuration.

Development environment	Tool and version number
The operating system	Ubuntu 16.04
Blockchain deployment container	Docker 19.03.9
Blockchain development tools	Hyperledger Fabric 1.4.0。
Hyperledger Fabric SDK Java 1.4.0	
The database	MySQL 5.7
System development language	Java1.8.0261
System backend development framework	Springboot 2.1.4, MyBatis 2.0.1
System front-end development framework	Vue.js2.0, Uniapp

enterprises, commercial enterprises, and end consumers. Data entry, data verification, and data maintenance are carried out through simple operation. Traces the process in the design, the most will include red jujube cultivation in the process of planting, management, and collect key information such as input to the database and the most maintenance blockchain node, after the consensus of the processes in the system to write the key information into blocks, once written, this information is given by the blockchain traceability, difficult to tamper with the traits, such as in the information written after success, will get a block return value, the return value will be saved in the database maintained by the market supervision administration, is the key to realize consumers on the date planting time, location, and other key information query. Collectively, red jujube acquisition time can be purchase level key information into blocks, such as industrial enterprises processing time, production time, and so on key information can be written to block, the business enterprise can be dispatching information, storage time, put in the region during the key information such as writing block, get the corresponding return value, save in the administration of the market supervision, and maintenance of the database. The market supervision administration plays a supervisory role in the whole process of data entry, providing national supervision for consumers to purchase dates and processed products, and effectively protecting the rights and interests of consumers. As far as consumers are concerned, detailed information about dates and processed products provided by jujube farmers, jujube gardens, industrial enterprises, and commercial enterprises can be easily found through the quality traceability system. Its main operating process is: the consumer by scanning red jujube and its processed products of QR code on the package to the market supervision and administration of the maintenance of red jujube-related data in the database query, can detailed process of red date information query, if the query shows information does not accord with red jujube actual information, to a large extent can be concluded that the purchase of jujube is safe, consumer can undertake inform against to market supervisory bureau or commercial enterprise, obtain corresponding compensation.

EXPERIMENTS AND RESULTS

Choose the high-quality characteristics in XRJ varieties to quality and safety traceability system demonstration experiment, on the basis of sufficient study, analysis of demand, study and development to establish traceability anticounterfeiting XRJ whole-industry chain management system, unified data standard system construction, software and hardware, and the reliable technology are constructed to obtain integrity, practicality, advanced, and scalability. Anticollision method of RFID tag in large dynamic change and fast moving environment is carried out synchronously and is used in traceability anticounterfeiting management system to improve batch tag identification efficiency in complex environment. The advantages and characteristics of the XRJ whole-industry chain traceability anticounterfeiting management system based on radiofrequency identification technology will be gradually promoted to red jujube, planting, and sales enterprises in the province, benefiting the majority of villages and growers.

According to the distribution of relevant personnel in the enterprise, the application scheme is determined: (1) personnel in the headquarters. Due to data center server deployment in enterprise headquarters, system administrator in the enterprise headquarters, production management personnel, and technical personnel can switch to the traditional C/S architecture system client data reporting and can carry out data analysis and statistics, the traditional C/S architecture system data reporting and management of the purpose is to improve the response speed data, reduce the influence of Windows Communication Foundation (WCF) frame on system performance; (2) Personnel other than the headquarters. The staff in the base, processing plant, or outside can choose different system clients for data reporting and management according to whether the network is smooth. When the network is not smooth, people can choose the traditional C/S frame system client to input the local data into the local server and carry out the local data analysis and statistics. When the network is normal, people can use the WCF client to synchronize data from the local server to the data center at the headquarters. After the synchronization, the WCF client can directly report data to the data center at the headquarters in real time and analyze and collect statistics over the internet.

The Henan Xinzheng Xinxing Jujube Industry Corporation Ltd. is chosen to validate the system. The network infrastructure of the enterprise is better. The enterprise has a provincial planting base of red dates: Nankou red jujube production center and 9 large-scale planting bases. Some bases have poor network infrastructure and the network cannot be guaranteed to be smooth all the time. According to the network infrastructure of the enterprise, it is very suitable to use the agricultural production traceability management system based on hybrid architecture. At present, the client of the whole system has been deployed in three planting bases and the server has been installed in the corporate headquarters. The production management department is in urgent need of obtaining the real planting area, number of trees planted, amount of pesticides and fertilizer applied and their types, yield, and other data of each base. However, such data can only be collected after the production cycle is completed

TABLE 2 | Test results of each module of the system.

Test module	Function	Input data	Expected result	Test result
Farming operations	Farming Operations	Agr. operation data	Save success	Normal
Quality inspection	Quality Inspection Data	Quality check data	Save success	Normal
Warehouse storage	Storage Repository Data	storage repository data	Save success	Normal
Transport	Shipping data	Transportation data	Save success	Normal
Data auditing	Farm process audit	Farm operation ID	The audit is successful and the data	Normal
	Warehouse storage audit	Storage ID		Normal
	Quality inspection audit	Quality test ID		Normal
	Transport data audit	Transport data ID		Normal
Traceability data tamper verification	Clothing operation tamper effect test	Farm operation ID	If not tampered, the content not displayed. Otherwise, the data comparison between the database and the blockchain is displayed	Normal
	Tamper check is stored in the library	Storage ID		Normal
	Quality test tamper check	Quality test ID		Normal
	Transportation data tamper check	Transport data ID		Normal
	Data reduction	Farm operation ID Quality test ID Storage ID Transport data ID	Data restoration succeeds	Normal
Product management	New product	Product information	New success	Normal
	Edit product	Product information	Modify success	Normal
	Remove product	Product ID	Delete success	Normal
	Query products by terms	Query conditions	Query products meet conditions	Normal
	Traceability queries	Product Batch No.	Data of origin, product, agricultural operation, warehousing, storage and transportation of this batch of products are queried from the blockchain	Normal

for a period of time. After adopting this system, the above data can be obtained on the same day. In addition, the following functions are also realized: (1) query the planting and production situation of different batches, different people, different bases, and different time periods; (2) Trace codes are formed according to relevant rules; (3) Query the quality management of the whole production process according to the traceability code; and (4) Upload enterprise traceability data to the ministerial traceability data center through the upload component. **Figure 3** shows the main interface of XRJ production traceability management system. The upper part is divided into the system menu, the left part is divided into the navigation bar, and the right part is divided into the main part of the system, by clicking the specific items in the navigation bar and people can operate the corresponding record table on the right.

The blockchain-based agricultural product traceability system consists of six modules: source and origin, product, agricultural operation, quality inspection, warehousing and storage, and transportation. The data of agricultural operation, quality inspection, warehousing and storage, and transportation shall be stored in the blockchain after verification. We collectively refer to the data of these four modules as source information. The specific implementation of each functional module of the traceability system is demonstrated through sequence diagram analysis. The system development environment configuration is shown in **Table 1**. Hyperledger Fabric network deployment requires the use of Docker container because the Fabric's billing nodes, sorting nodes, and smart contracts all run in Docker container. In system terminal sudo, add apt—repository "deb [arch = amd64] https://hyperledger-fabric.readthedocs.io/en/release-1.2/network/network.html stable" and Sudo apt-getinstall docker-ce docker-ce-cli container. IO installs the Docker container.After setting up the Fabric operating environment, the next step is to build the Fabric blockchain, consisting of the following major steps: generating certificates and keys, generating genesis blocks, creating channels, starting the Fabric network, nodes joining channels, and installing and instantiating smart contracts.

In the development stage of the system, white-box test is applied to simulate all the kinds of possibilities inside the program to improve the robustness of the processing method. In the test stage, black-box test is used to test each functional module by simulating user behavior. The test results of each module of the system are shown in **Table 2**.Each functional module of the system is tested by building a Fabric blockchain and deploying smart contracts. As shown in **Table 2**, the system completes the development and implementation of specific functions to ensure the reliability of the system.

CONCLUSION

With the rapid development of economic level, some agricultural product enterprises reduce production costs to maximize the profits of some agricultural enterprises and result in a lot of agricultural product safety incidents in recent years. Therefore, people have to pay more attention to food safety in daily life. This article constructs a set of XRJ quality traceability system and carries on the system analysis and application verification. The hybrid model in the system makes use of its unique time stamp, consensus mechanism, and other technical means to realize the data of jujube planting, jujube purchasing, jujube

processing, jujube sales, and other data that cannot be tampered and traceable. At the same time, the market supervision bureau and consumers are included in the date quality traceability supervision system, which breaks the information island of traditional traceability, provides information support, realizes the full transparency of production and sales process to a certain extent, provides technical support for good market operation order, and provides quality guarantee for consumers' life and health. With in-depth study and development of blockchain technology, coupled with the requirements of high-quality development of the date industry, blockchain technology will provide new solutions to date traceability, pesticide residues, and brand counterfeiting. While establishing XRJ safety supervision and supervision mechanism, relevant departments need to establish traceability XRJ system, implement relevant responsibility to individuals, and strengthen food production supervision at the source.

DATA AVAILABILITY STATEMENT

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

ETHICS STATEMENT

Written informed consent was obtained from the individual(s) for the publication of any identifiable images or data included in this article.

AUTHOR CONTRIBUTIONS

RJ contributed to the model construction and manuscript writing. PL contributed to the data acquisition and experimental verification.

FUNDING

This study was supported by the Key Science and Technology Project of Science and Technology Department of Henan Province (Nos. 212102210404, 222102110134, 222102110280, and 222102210122), the Science and Technology Project of Education Department of Henan Province (Nos. 22B520049 and 21B520024), and the 14th Five-year Plan of Education Science of Henan Province (No. 2021YB0499).

REFERENCES

Almalki, F. A., Soufiene, B. O., Alsamhi, S. H., and Sakli, H. (2021). A Low-Cost Platform for Environmental Smart Farming Monitoring System Based on IoT and UAVs[J]. *Sustainability* 13, 1–26. doi: 10.3390/su13115908

Alsamhi, S. H., Almalki, F. A., Afghah, F., Hawbani, A., Shvetsov, A. V., and Lee, B. (2022). Drones' Edge Intelligence Over Smart Environments in B5G: blockchain and Federated Learning Synergy. *IEEE Trans. Green Commun. Netw.* 6, 295–312. doi: 10.1109/TGCN.2021.3132561

Alsamhi, S. H., Lee, B., Guizani, M., Kumar, N., Qiao, Y., and Liu, X. (2021). Blockchain for decentralized multi-drone to combat COVID-19 and future pandemics: framework and proposed solutions. *Trans. Emerg. Telecommun. Technol.* 32:e4255. doi: 10.1002/ett.4255

Aung, M. M., and Chang, Y. S. (2014). Traceability in a food supply chain: safety and quality perspectives. *Food Control* 39, 172–184. doi: 10.1016/j.foodcont.2013.11.007

Christidis, K., and Devetsikiotis, M. (2016). Blockchains and Smart Contracts for the Internet of Things. *IEEE Access* 4, 2292–2303. doi: 10.1109/ACCESS.2016.2566339

Cocco, L., and Mannaro, K. (2021). "Blockchain in Agri-Food Traceability Systems: a Model Proposal for a Typical Italian Food Product," in *IEEE International Conference on Software Analysis, Evolution and Reengineering (SANER)*, (Honolulu, HI, USA: IEEE). doi: 10.1109/SANER50967.2021.00085

Corallo, A., Latino, M. E., Menegoli, M., and Striani, F. (2020). The awareness assessment of the Italian agri-food industry regarding food traceability systems. *Trends. Food Sci. Technol.* 101, 28–37. doi: 10.1016/j.tifs.2020.04.022

Demestichas, K., Peppes, N., Alexakis, T., and Adamopoulou, E. (2020). Blockchain in Agriculture Traceability Systems: a Review. *Appl. Sci.* 10:4113. doi: 10.3390/app10124113

Feng, H., Wang, X., Duan, Y., Zhang, J., and Zhanga, X. (2020). Applying blockchain technology to improve agri-food traceability: a review of development methods, benefits and challenges. *J. Clean. Prod.* 260:121031. doi: 10.1016/j.jclepro.2020.121031

Galvez, J. F., Mejuto, J. C., and Simal-Gandara, J. (2018). Future challenges on the use of blockchain for food traceability analysis. *TrAC Trends Anal. Chem.* 107, 222–232. doi: 10.1016/j.trac.2018.08.011

Konstantinidis, I., Siaminos, G., Timplalexis, C., Zervas, P., Peristeras, V., and Decker, S. (2018). "Blockchain for Business Applications: A Systematic Literature Review," in *International Conference on Business Information Systems*, eds W. Abramowicz and A. Paschke (Cham: Springer), 384–399. doi: 10.1007/978-3-319-93931-5_28

Mukkamala, R. R., Vatrapu, R., Ray, P. K., Sengupta, G., and Halder, S. (2018). Blockchain for Social Business: Principles and Applications[J]. *IEEE Eng. Manag. Rev.* 46, 94–99. doi: 10.1109/EMR.2018.2881149

Rakitskii, V. N., Doan, N. H., Fedorova, N. E., and Bereznyak, I. V. (2020). Safety of imported agricultural products: pesticide residues. *Zdr. Ross. Fed. Minist. Zdravookhraneniia RSFSR* 64, 150–157. doi: 10.46563/0044-197X-2020-64-3-150-157

Ray, P., Duraipandian, R., and Sinha, R. A. (2021). Framework for Identifying and Managing Risk Impact Factors for Disruptions in the Food Supply Chain. *Int. J. Eng. Appl. Sci.* 12, 1–11.

Tharatipyakul, A., and Pongnumkul, S. (2021). User Interface of Blockchain-Based Agri-Food Traceability Applications: A Review. *Access IEEE* 9, 82909–82929. doi: 10.1109/ACCESS.2021.3085982

Yadav, S., Luthra, S., and Garg, D. (2020). Internet of things (IoT) based coordination system in Agri-food supply chain: development of an efficient framework using DEMATEL-ISM. *Oper. Manag. Res.* 1–27. doi: 10.1007/s12063-020-00164-x

Yang, C., and Sun, Z. (2020). "Data Management System based on Blockchain Technology for Agricultural Supply Chain," in *International Conference on Data Mining Workshops (ICDMW)*, (Sorrento: IEEE). doi: 10.3390/s20102990

Zhang, X., Cao, Z., and Dong, W. (2020). Overview of Edge Computing in the Agricultural Internet of Things: Key Technologies, Applications, Challenges. *IEEE Access* 8, 141748–141761. doi: 10.1109/ACCESS.2020.3013005

Zhang, Y., Li, J., and Ge, L. (2021). "Research on Agricultural Product Supply Chain Based on Internet of Things and Blockchain Technology," in *Cyber Security Intelligence and Analytics. CSIA 2021. Advances in Intelligent Systems and Computing*, eds Z. Xu, R. M. Parizi, O. Loyola-González, and X. Zhang (Cham: Springer). doi: 10.3390/s22041304

Zhao, G., Liu, S., Lopez, C., Lu, H., Elgueta, S., Boshkoska, B. M., et al. (2019). Blockchain technology in agri-food value chain management: a synthesis of applications, challenges and future research directions. *Comput. Ind.* 109, 83–99. doi: 10.1016/j.compind.2019.04.002

Zhao, Y., and Ning, C. (2017). "Research on Traceability of Agricultural Products Based on Internet of Things," in *IEEE International Conference on Computational Science & Engineering*, (Piscataway: IEEE). doi: 10.1109/CSE-EUC.2017.264

Zu, X., He, Y., Pu, Y., and Yang, L. (2020). "The Characteristics of Agricultural Product Quality and Safety Crisis Based on Content Analysis Method," in *14th International Conference on Management Science and Engineering Management*, eds J. Xu, G. Duca, S. Ahmed, F. García Márquez, and A. Hajiyev (Cham: Springer), 119–129. doi: 10.1007/978-3-030-49829-0_9

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

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

frontiers | Frontiers in Plant Science

ORIGINAL RESEARCH
published: 09 June 2022
doi: 10.3389/fpls.2022.864458

Fast Location and Recognition of Green Apple Based on RGB-D Image

Meili Sun[1], Liancheng Xu[1], Rong Luo[2], Yuqi Lu[1] and Weikuan Jia[1,3]**

[1] School of Information Science and Engineering, Shandong Normal University, Jinan, China, [2] State Key Laboratory of Biobased Materials and Green Papermaking, Qilu University of Technology (Shandong Academy of Sciences), Jinan, China, [3] Key Laboratory of Facility Agriculture Measurement and Control Technology and Equipment of Machinery Industry, Zhenjiang, China

In the process of green apple harvesting or yield estimation, affected by the factors, such as fruit color, light, and orchard environment, the accurate recognition and fast location of the target fruit brings tremendous challenges to the vision system. In this article, we improve a density peak cluster segmentation algorithm for RGB images with the help of a gradient field of depth images to locate and recognize target fruit. Specifically, the image depth information is adopted to analyze the gradient field of the target image. The vorticity center and two-dimensional plane projection are constructed to realize the accurate center location. Next, an optimized density peak clustering algorithm is applied to segment the target image, where a kernel density estimation is utilized to optimize the segmentation algorithm, and a double sort algorithm is applied to efficiently obtain the accurate segmentation area of the target image. Finally, the segmentation area with the circle center is the target fruit area, and the maximum value method is employed to determine the radius. The above two results are merged to achieve the contour fitting of the target fruits. The novel method is designed without iteration, classifier, and several samples, which has greatly improved operating efficiency. The experimental results show that the presented method significantly improves accuracy and efficiency. Meanwhile, this new method deserves further promotion.

Keywords: green target fruit, center location, density peak clustering, kernel density estimation, RGB-D image

OPEN ACCESS

Edited by:
Lei Shu,
Nanjing Agricultural University, China

Reviewed by:
Abbas Atefi,
California Polytechnic State University, United States
Yunchao Tang,
Guangxi University, China

***Correspondence:**
Liancheng Xu
lchxu@163.com
Weikuan Jia
jwk_1982@163.com

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

Received: *28 January 2022*
Accepted: *26 April 2022*
Published: *09 June 2022*

Citation:
Sun M, Xu L, Luo R, Lu Y and Jia W (2022) Fast Location and Recognition of Green Apple Based on RGB-D Image. Front. Plant Sci. 13:864458. doi: 10.3389/fpls.2022.864458

INTRODUCTION

The machine vision system, employed to realize the localization and recognition of target fruits, has been widely applied in agricultural production processes, such as apple orchard yield estimation (Maheswari et al., 2021), apple automatic harvesting (Jia et al., 2020a), and fruit growth monitoring (Genno and Kobayashi, 2019). How to achieve accurate recognition and fast location of target fruits will directly affect the reliability and real-time performance of automated operations. This is also the key to research, which has attracted the attention of many domestic and foreign scholars. In the recent years, whether it is with the help of the monocular vision system (Zhang et al., 2017) or binocular vision system (Si et al., 2015), a single fruit (Silwal et al., 2017) or overlapping occluded fruit (Lv et al., 2019b), a static target (Häni et al., 2020a), or a dynamic target (He et al., 2017), it has made great progress in recognition, and most of these studies are implemented with red apples as the harvesting target. However, many types of apples are green apples, such as white winter Pearmain and Granny Smith. Because the color of the fruit is similar to the color of the background branches and leaves, it is difficult to recognize and locate the target fruit. Besides, affected by factors,

such as complicated unstructured orchard environment and the blurred borders of irregular bright areas, the performance of visual system is relatively not good.

It is very necessary to improve the recognition and location efficiency of green apples and further improve the automated assembly level of the vision system in agricultural production management. Accurate location and fast recognition of green target fruit become a new challenge. With the joint efforts of many scholars, certain progress has been made (Lv et al., 2019a; Sun et al., 2019; Behera et al., 2020; Ji et al., 2020). It is difficult to recognize green apples only from the perspective of color, and it needs to be processed or try other features. Therefore, Zhang used a color camera equipped with a ring flash to capture images of apple trees. By analyzing the feature difference between the green apple areas and the similar-color background, a classifier was designed based on color features to recognize the green apples in the near-color background. The recognition success rate reached 89.3%, and the algorithm took 3.14s on average (Zhang et al., 2014).

To recognize the same color fruits under unstable light and the occlusion of branches and leaves, Barnea proposed a 3D detection model based on RGB color features and related shape features (Barnea et al., 2016), where the mean average precision of recognition reached 55%. Li DH et al. presented an improved spectral clustering algorithm based on mean shift and sparse matrix to recognize overlapping green apples and the algorithm improved the running speed (Li et al., 2019), The high coincidence degree of the optimized spectral clustering algorithm reached 95.41%, and the false detection rate is 3.05%. Gaussian linear fitting on the foreground images of the green apples was adopted under the V channel, and threshold segmentation was used to segment the images (Li et al., 2018), which was 91.84% of the recognition rate. The gPb-OWT-UCM edge detection algorithm was applied in the green apple detection model based on the SUN saliency detection mechanism by Wang, which can obtain complete and accurate green apple segmentation images (Wang et al., 2019). The above method reached accurate segmentation with average sensitivity, false-positive rate, false-negative rate, and overlapped rate of 8.4, 0.8, 7.5, and 90.5%, respectively. The segmentation time of each apple image was around 37.1 s. All these methods merged corresponding texture features through color features. Due to the changes in light intensity and angle, the boundary between the target fruit and the background is blurred, and the recognition effect is not ideal. Affected by factors, such as occlusion of branches and leaves and overlapping between fruits, some features are absent, which makes the recognition difficult. Some algorithms have high time and space complexity, and it is difficult to meet real-time operation requirements.

With the rapid development of deep learning theory, end-to-end automatic detection process based on deep networks and the advantages of deep extraction of image features (Kamilaris and Prenafeta-Boldú, 2018; Koirala et al., 2019; Zhao et al., 2019; Tang et al., 2020) bring a new perspective for recognizing green apples. A multi-scale multilayer perceptron and a convolutional neural network (CNN) were applied to segment apple images and extract the apple target by Bargoti, where the apple target was recognized and counted by the watershed segmentation and the circular Hough transform method (Bargoti and Underwood, 2017), where F1-score reached 85.8%. To accurately locate the tomato fruit under the complex scenes, Liu improved the YOLOv3 one-stage target detection model to predict the circular areas (Liu et al., 2020), and the precision and times of YOLO-Tomato were 94.75% and 54 ms. Jia optimized the Mask R-CNN to adapt to the detection of apple targets, where the residual neural network (ResNet) and dense convolutional network (DenseNet) were combined as the feature extraction network of the original model that the precision and recall rate have reached 97.31 and 95.70%, respectively (Jia et al., 2020b). Li proposed an ensemble U-Net segmentation model, and the high-level semantic features of U-Net and the edge features of Edge were integrated to retain multi-scale contextual information and realize efficient segmentation of target fruit (Li et al., 2021), where the recognition rate reached 95.11% and the recognition speed was 0.39 s. A modified YOLOv3 model based on clustering optimization is designed, and the influence of front-lighting and backlighting is clarified to detect and recognize banana fruits, inflorescence axes, and flower buds by Wu et al. (2021). To recognize and detect plant diseases, Chen et al. (2022) proposed an improved plant disease-recognition model based on the YOLOv5 network model *via* a new involution bottleneck module, an SE module, and an efficient intersection over union loss function to optimize the performance of target detection, where mean average precision reached 70%. A DaSNet-V2 network structure was proposed by Kang, visual sensors were used for real-time detection and instance segmentation of orchard apples, and the branches were segmented. Visual sensors were applied in field trials, and the experimental results showed that the presented method was efficient and reliable (Kang and Chen, 2020), where the precision of detection, fruit segmentation, and branch segmentation achieved 0.844, 0.858, and 0.795, and computational time was 30 and 265 ms on GTX-1080Ti and Jetson-TX2, respectively. To improve the performance of apple detection, a deep learning method approach based on the adaptive training sample selection (ATSS) was applied to close-range and low-cost terrestrial RGB images (Biffi et al., 2020). In addition, considering the lack of public datasets in the field of fruit detection, the MinneApple benchmark dataset (Häni et al., 2020b) and KFuji RGB-DS dataset (Gené-Mola et al., 2019) are publicly used to study fruit detection and segmentation. The MinneApple benchmark dataset is designed and published for detection and segmentation. The KFuji RGB-DS dataset is presented, containing 967 multi-modal images and 12,839 Fuji apples. Although the recognition accuracy of the target fruit is high based on deep learning, it has high requirements for machine hardware and requires a large number of training samples. In the actual operation process, it is difficult to meet these conditions one by one.

For the above problems, with the help of gradient information of depth images, an improved density peak cluster segmentation method of RGB images is proposed to fast location and accurate recognition of green apples based on RGB-D image information. The depth image of the apple is used to obtain the center of the target fruit and determine the location of the target fruit.

To obtain the target fruit area, an optimized density clustering segmentation algorithm is introduced to segment the RGB image of the apple. Next, the depth and RGB images are combined by scanning the maximum radius of the segmented area where the circle center is located, and the target fruit contour is fitted to realize the efficient recognition and location of green apples. The new method applies the process of locating the circle center first and then recognizing, breaking the traditional method of recognizing first and then locating. Meanwhile, the new method is designed without iteration and a classifier, which effectively improves the efficiency of recognizing and locating the target fruit. In addition, the new method can complete the algorithm training for small sample datasets, greatly saving sample processing time. To summarize, our contributions to this article are as follows.

(1) The depth image information is applied to locate the fruit center *via* the idea of vorticity.
(2) An optimized density peak cluster segmentation method based on kernel estimation is designed to segment RGB images without subjective judgment.
(3) Experimental results outperform the other state-of-the-art models in accuracy and efficiency for green fruit recognition and detection.

The rest of this article is organized as follows. In section "Materials and Methods," image acquisition and ideas of this article are introduced. In section "The Center Location of the Target Fruit," we describe in detail how to use depth information to locate the target fruit. Section "Target Area Segmentation" illuminates how to apply color image information to segment the target. Section "Fitting of Target Fruit" is to fit the target fruit according to the fast-locating result of the depth image and the segmentation result of the color information. In section "Results and Discussion", the experiments, including the experimental design and result analysis, are conducted and experimental results can be discussed. The conclusion is presented in section "Conclusion."

MATERIALS AND METHODS

Image Acquisition

Image acquisition location: Longwangshan apple production base in the Fushan District, Yantai City, Shandong Province (the agricultural information technology experimental base of Shandong Normal University). All images were collected under natural light.

Image acquisition device: Kinect V2 (Microsoft). The RGB images have a resolution of 1,920 by 1,080 pixels whereas the depth images have a resolution of 512 by 424 pixels. All RGB images and depth images were stored in the bitmap (BMP) format. The distance range from the camera lens to the target is 0.5–4.5 m. When the Kinect camera acquired the depth image, it collected each point in the field of view and forms a depth image representing the surrounding environment.

Apple category: Gala. Fruits were in an immature status before harvesting.

Image acquisition mode: To imitate the actual monitoring environment, a tripod was used to fix the Kinect V2 camera angle. Kinect SDK 2.0 software was applied to capture the RGB-D images from the same angle. The color and depth images were saved in Portable Network Graphics (PNG) format. For the same fruits, we collected the images from different angles, including a single target, branch and leaf occlusion, or overlapped fruits. To facilitate image fusion, the resolution of all RGB images was adjusted to be the same as the resolution of the depth image 512 by 424 pixels. The images of green apples in actual environments, such as single fruits, overlapping fruits, and occlusion fruits, are shown in **Figure 1**.

The Method and Ideas of This Article

Based on the features of green apples, this article proposes a method for fast location and recognition of the green target fruit by the depth images (Choi et al., 2017) and RGB images of the target fruit. The depth image is utilized to find the center location of the target fruit and realize the effective location of the target fruits. Then, an RGB image is applied to segment fruits and background and obtain the target fruit area. Finally, the depth images and the RGB images are merged into a segmented area with the circle center as the target fruit, realizing the location and recognition of the target fruit. The overview of the new method is shown in **Figure 2**. The new method has two branches: locating the center of the target fruit through the depth image; segmenting the target fruit area from the RGB image.

For the depth image, the depth information of the target image is analyzed to draw an iso-contour map of the depth image and obtain the gradient field information of the target image. The three-dimensional gradient information of the target fruit is projected to the two-dimensional. To achieve the accurate location of the target fruit, the two-dimensional gradient vector is rotated 90° in the clockwise direction. It found that vorticities of the target fruit are relatively regular and orderly, whereas background regions present divergent and irregular states. According to the geometric meaning of the depth image, the center of each vorticity is viewed as the center of the target fruit.

A density peak clustering segmentation algorithm based on kernel density estimation is proposed in RGB images. First, a kernel density estimation method is applied to calculate the local density and difference in each data point in a non-parametric way. Note that a super-pixel segmentation algorithm is applied to segment images into blocks to convert images from pixel-level representation to area-level representation, which can reduce the number of data points and the amount of calculation. Then, a decision graph is constructed to separate the cluster centers and other data points. Next, a double sorting algorithm is adopted to automatically find the cluster centers, achieving the green target fruit segmentation.

Based on the center of the target fruit and the target fruit segmented area, the center of the target fruit is merged with the segmented area. Specifically, the segmented areas located in the circle center are considered as the green target fruits; the maximum value method is employed to obtain the radius of the target fruit. Finally, the contour of the target fruit is fitted to

realize the location and recognition of the target fruit through the circle centers and its corresponding radius.

THE CENTER LOCATION OF THE TARGET FRUIT

Traditional target fruit location is mostly to segment the target fruit area first and then find the circle center through morphological methods. The location, detection, and segmentation results mentioned by these methods in the Introduction section are not accurate enough and the calculation amount is relatively large. In complex unstructured orchard environments, the accurate location of the target fruit is still an unsolved problem, and the locating of the green target fruit is even more challenging. This study introduces the depth information of the target fruits and applies the spatial information of the images to find the centers of the green apples.

Contour Image Acquisition

The Kinect camera can directly obtain the distance information between the target and lens. Based on the imaging principle of the depth sensor, the pixel closer to the camera has a smaller distance value. The smaller distance value of the target corresponds to the smaller pixel value. The larger distance value of the surrounding background corresponds to the larger pixel value. Therefore, the pixels of the same value in the depth image are fitted to a closed curve. These curves can be projected onto a two-dimensional plane. Finally, the depth contour of the target image is obtained, that is, the depth of the pixels on the same depth contour is equal.

For the iso-contour map, a simple smoothing filtering process is performed. The three-dimensional geometric characteristics of the target are calculated through the mapped depth information. The depth of the image is drawn with a certain depth difference. **Figure 1A** will be an example to elaborate on the fast location and accurate recognition of green apples, as shown in **Figure 3**.

Depth Image Gradient Information

According to the principle of the depth sensor, the smallest pixel from the objects to the camera has the smallest distance value. As shown in **Figure 3**, the depth value of the central area of the target fruit is less than that of the non-central area of the target fruit based on the numerical characteristics of the depth image. Therefore, the depth information of the target depth image can be quantified by the gradient field theory. A vector $\overrightarrow{V}$ can be constructed.

$$\overrightarrow{V} = \nabla D = \left(\frac{\partial D}{\partial x}, \frac{\partial D}{\partial y}, \frac{\partial D}{\partial z}\right) = (u, v, w) \tag{1}$$

where $\overrightarrow{V}$ represents a set of vectors (u, v, w) on the gradient vector field in three-dimensional space; u, v, w are partial derivatives of depth D in the x, y, z directions in three-dimensional coordinates, respectively, obtaining the gradient direction of the target fruits. Since the surface of the target fruit is convex, the 3D gradient field information can be projected onto the 2D plane. In other words, only the x and y directions are considered.

$$\overrightarrow{V} = \nabla D = \left(\frac{\partial D}{\partial x}, \frac{\partial D}{\partial y}\right) = (u, v) \tag{2}$$

where u, v represent the partial derivatives of the gradient field D in the x, y directions, respectively, which can obtain the gradient direction of the target fruit. Then, the direction vector on the 2D plane is obtained, which diverges outward along the apple surface in the gradient field, as shown in **Figure 4**.

In **Figure 4**, it can be found that the obtained direction vector presents a state of outward divergence in a relatively orderly way for the relatively regular target fruit area. Due to the complex

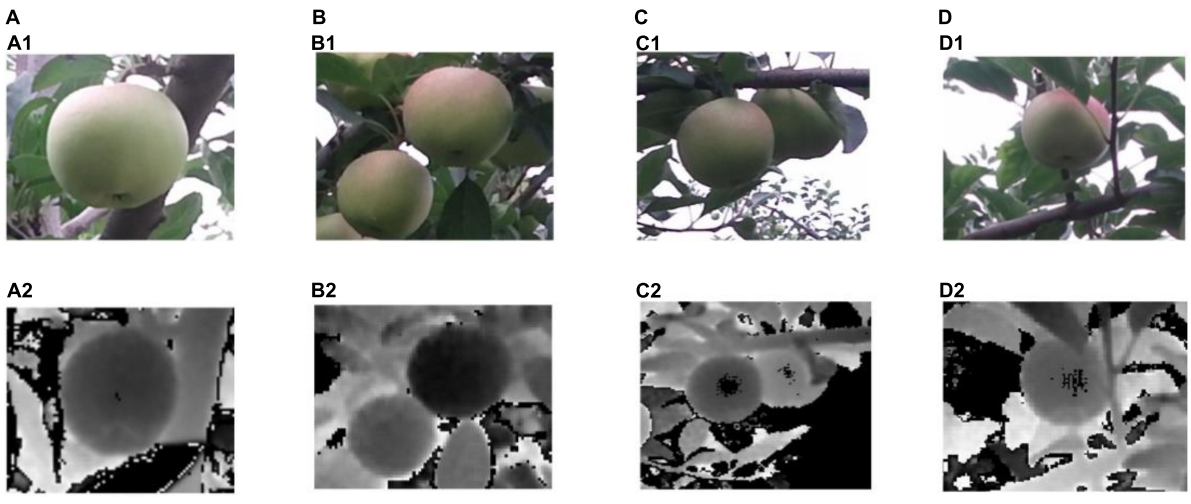


FIGURE 1 | Apple images of RGB and depth images collected at the same sample point (note: panels **(A,B)** are single fruits without obscured; **(C)** are overlapping fruits; **(D)** are occlusion fruits; top images of every group represent the RGB image and bottom images represent depth images). The number of images: This study collected a total of 300 sample points data, including 300 depth images and 300 RGB images and containing 526 target fruits in the dataset.

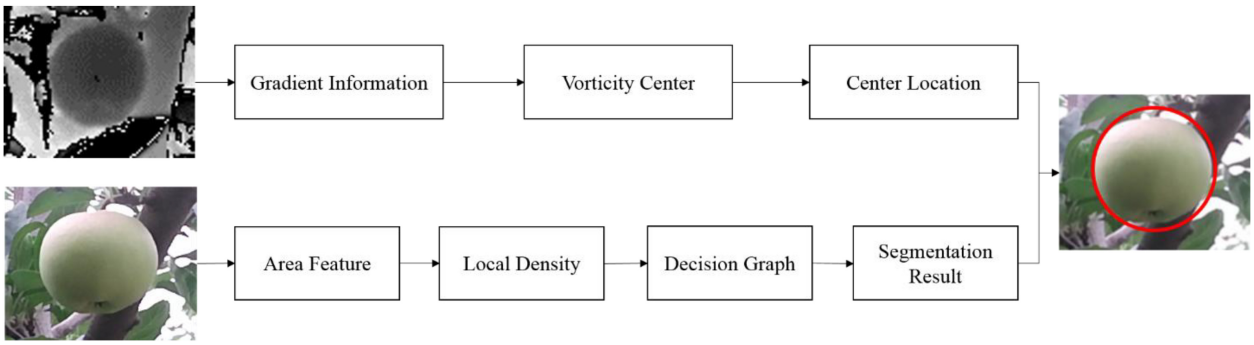


FIGURE 2 | The overview of our method.

background and the slightly disordered depth information, the vector projection is more chaotic in the non-target fruit area. In other words, the more regular the target fruit is, the more orderly the vector projection is. The more complex the background is, the more disorderly the vectors are.

Center Location

Enlighted by dynamic vorticity, we apply the vorticity center of the target depth image as the fruit centers. Specifically, the vorticity is a 3D vector, and a pixel angular velocity (only direction) is introduced through the adjacent vectors.

$$\vec{\omega'} = \nabla \times \vec{V} = \left(\frac{\partial}{\partial x}, \frac{\partial}{\partial y}, \frac{\partial}{\partial z}\right) \times (u, v, 0) = \left(\frac{\partial u}{\partial x} - \frac{\partial v}{\partial y}\right) \vec{z} \tag{3}$$

where ∇ is a partial derivative operator in x, y, z directions, and $\vec{z}$ represents a direction vector on the z-axis.

Since the rate of change of the vortex size increases from the edge to the center of the circle, the vectors in the gradient field are spliced into a similar arc shape. If the vector size is the same, the angular velocity is defined as follows:

$$\omega = \frac{\phi}{t} \tag{4}$$

where ϕ represents the size of the arc length (similar to the size of the vector), and t represents the same time.

When gradient vectors with convex parabola are rotated in the same direction, a similar convergent vortex is formed in the position of vorticity maximum value and fruit center. In other words, the gradient vector of the primitive concave paraboloid is rotated 90° in a clockwise direction, and it becomes divergent and disordered. The maximum vorticity of the fruit area can be calculated at the center point of the fastest distance change, namely, the center of the target fruit. Further, the gradient vector of the fruit center area is a convex parabola center with such a characteristic: the vorticity of the central region of the fruit is greater than that of the surrounding region. In other words, the vorticity value of the fruit gradually decreases from the fruit center to the fruits boundary. Therefore, this geometric characteristic causes the gradient vector field plot of apples in 2D appeared to stretch out from the center of the apples to the direction of its closest perimeter pixel. The center of the vorticity can be projected onto a 2D plane, and the gradient vector is rotated in the same direction along the depth contour. The central location is the center of the target fruit, as shown in **Figure 5**.

It can be found that the vorticity centers of the regular area are the centers of the target fruits after rotation of the gradient vector in **Figure 5**. This method can rapidly locate the centers of the target fruits.

TARGET AREA SEGMENTATION

For green apple image segmentation, it is difficult to achieve accurate segmentation by traditional image segmentation algorithms that only rely on color features due to the closeness of color between fruit skin and branches or leaves in the background. In this article, a density peak clustering algorithm is introduced to segment fruit area and background. To obtain a clearer segmentation boundary without human interference,

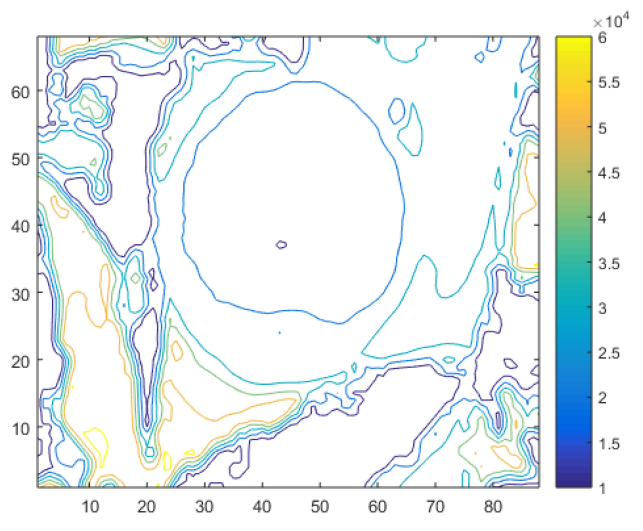


FIGURE 3 | Contours of the depth image.

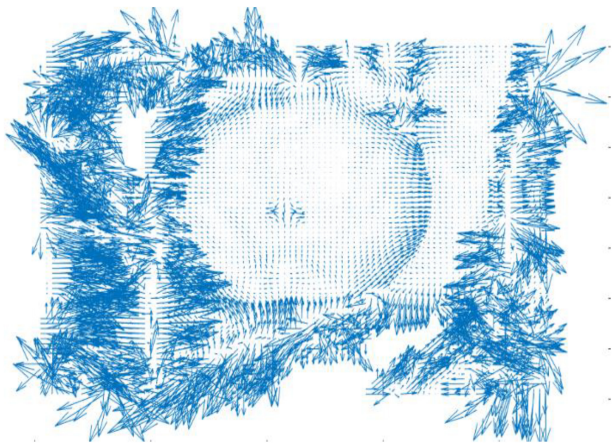


FIGURE 4 | Two-dimensional plane projection of gradient field.

a kernel density estimation method is utilized to optimize the clustering segmentation algorithm.

The main idea of the density peaks clustering algorithm (DPCA) (Rodriguez and Laio, 2014) is to judge the class that it belongs to by the neighboring points. The density of each class center point is greater than the density of neighboring points and this class center point is far enough away from the center points of other classes. Therefore, two main parameters are involved in the DPCA: the local density of data points and the distance between high-density points.

Image Block Segmentation

The basis of the clustering algorithm is to calculate the distance between data points, which can be transformed into the description of the feature components of the data points, and the feature components determine the distribution relationship of the data points in the feature space. In the field of image segmentation, the pixel points of the image are the data points in the clustering algorithm. However, pixels considered as the calculation unit consume a lot of calculation space and time. Therefore, an image is divided into blocks, and the image blocks are applied as data points for clustering segmentation, which can greatly reduce the amount of calculation. The most common method of image block division is super-pixel segmentation, which uses the simple linear iterative clustering (SLIC) (Achanta et al., 2012) super-pixel segmentation algorithm for gathering similar pixels in a small area to form irregular blocks.

Considering the small color difference between the target fruit and the background, we convert RGB color space into the L*a*b color space emphasizing color change before using the SLIC algorithm. The cluster center is initialized to C_k, and the iteration step is initialized to S. S can be formulated as follows:

$$S = \sqrt{\frac{N}{k}} \tag{5}$$

where N is the number of image pixels and k is the number of blocks. To prevent the cluster center from falling on the edge of the image, the gradient value of each pixel is calculated in the 3 by 3 neighborhood of the cluster center, and the center is moved to the pixel with the smallest gradient. Next, iterative optimization is performed. To save computing time, the color and spatial distance are calculated, where the distance Dis between each pixel and the cluster center is calculated the $2S$ by $2S$ the neighborhood around the center point.

$$\begin{cases} d_c = \sqrt{(l_i - l_j)^2 + (a_i - a_j)^2 + (b_i - b_j)^2} \\ d_s = \sqrt{(x_i - x_j)^2 + (y_i - y_j)^2} \end{cases} \tag{6}$$

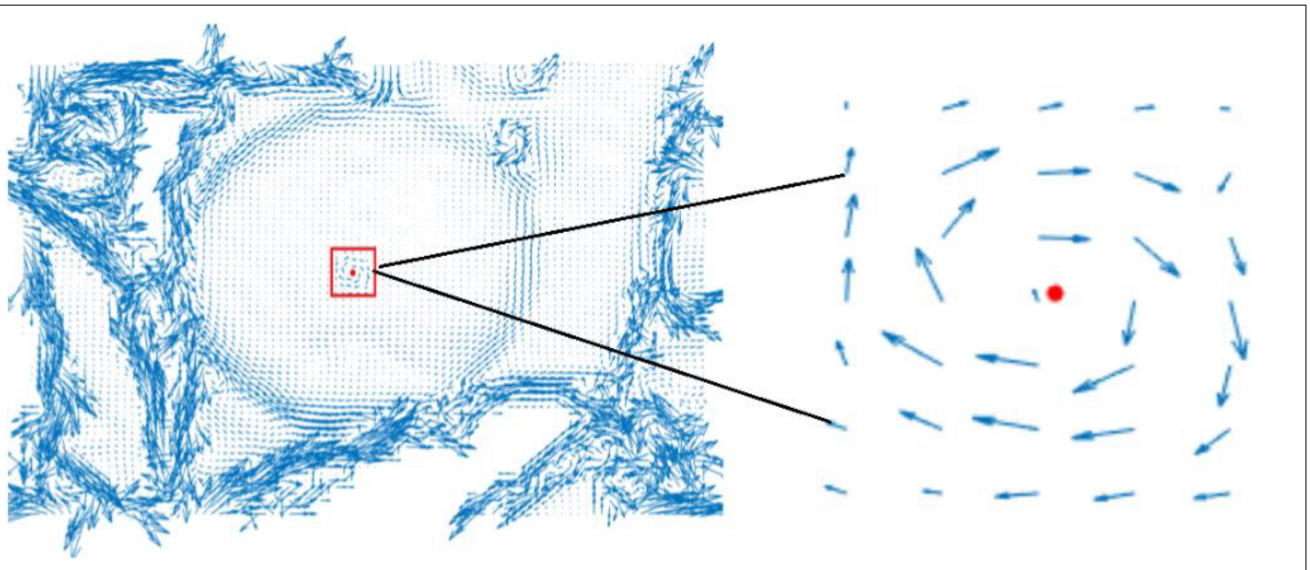


FIGURE 5 | Location of target fruits circle center (gradient vectors are uniformly rotated 90° clockwise).

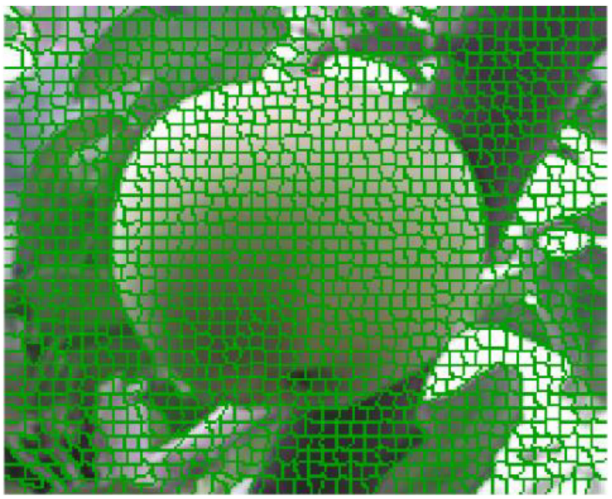


FIGURE 6 | SLIC super-pixel segmentation effect.

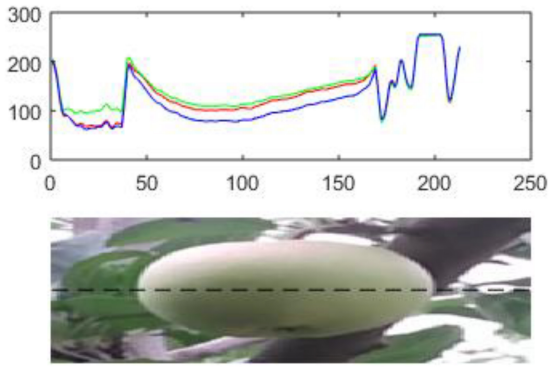


FIGURE 7 | The RGB components corresponding to different scenes in the image (note: the red line represents the R component; the green line represents the G component; the blue line is the B component; the three lines are the RGB components corresponding to the pixels in the black horizontal dashed line).

The distance *Dis* between each pixel and cluster center is constructed as follows:

$$Dis = \sqrt{(d_c)^2 + \left(\frac{d_s}{S}\right)^2 m^2} \tag{7}$$

where l, a, b are Lab space values, x, y are pixel coordinates, *and* m is the maximum possible value of Lab space distance. Each pixel updates the image block to which it belongs. The cluster center is updated by averaging the pixels of the same block until the center point no longer moves.

The SLIC image block segmentation algorithm considers the color and space features of the image and defines the search range of 2*S* by 2*S* effectively reducing the computation complexity of the DPCA. Finally, the target image block effect map is presented, as shown in **Figure 6**.

It can be seen from **Figure 6** that the image block is applied to represent the green apple image, which maintains the consistency of the original pixel points. These irregular blocks do not destroy the original image representation structure. Meanwhile, the target image is represented from the pixel level to the block level, which indicates that the conversion loss is small. Each irregular image block replaces pixels with the basic unit of the DPCA, which will greatly reduce the number of data points.

Feature Component Selection

After the target image is divided into image blocks, the effect of its shape and texture features is greatly reduced. Therefore, the feature component of the block is constructed in the unit of the super-pixel feature block as the feature components of the data points. Considering the closeness of color between the target fruit skin and the background, the color information of the target image is decomposed in the RGB color space to analyze features from different color channels to better select proper feature components, as shown in **Figure 7**.

It can be observed that the target fruit area in the green apple image has significant performance in the R and G channels, but is not obvious in the B channel in **Figure 7**. For the G channel, it is difficult to distinguish branches and leaves from the background. The performance of the B channel is not obvious for background branches and leaves and the target apple loses its significance. Therefore, R-B and G-B super-pixels are used to represent the image block features of the target image segmentation. Based on the image block, the R-B and G-B mean values are applied as the feature components of the data points to construct the feature space of the green apple target image through the density clustering segmentation algorithm, as shown in **Figure 8**.

The different colors of the data points represent the three classes obtained after clustering in **Figure 8**, where red, green, and yellow points represent green fruits, leaves, and branches, respectively. The boundary between the different clusters is not obvious and the shape is irregular. The data points of the boundary area are relatively scattered. Traditional clustering algorithms are difficult to deal with clustering problems with fuzzy boundaries, and low-density area boundaries are only presented inside the feature space.

Density Peak Clustering Segmentation Algorithm

The DPCA algorithm can find out the high-density areas segmented by low-density areas and obtain arbitrary-shaped clusters. The cluster centers are fixed, where the data points belong to the clusters with the closest distance and greater local density. DPCA meets two assumptions: the local density of the cluster center (density peak point) is greater than the density of surrounding data points; the distance between different cluster centers is relatively far. To find the class center that meets the two conditions at the same time, the local density ρ_i of each data point x_i and the distance δ_i from each data point x_i to the data point x_j are calculated, where the local density of the data

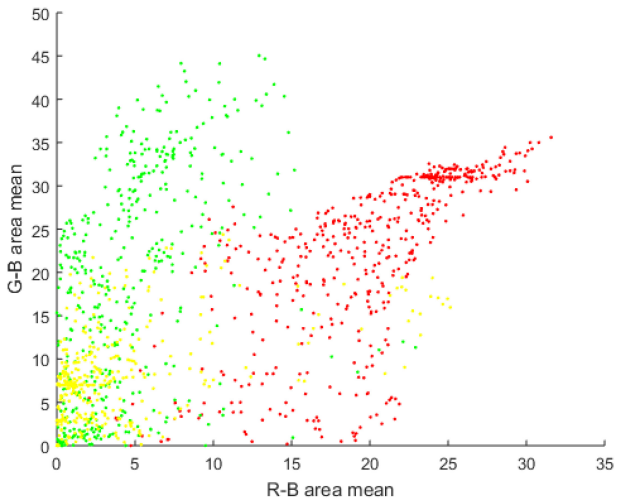


FIGURE 8 | The feature space of the mean value of R-B and G-B.

point x_j is greater than that local density of data point x_i and the distance is the closest.

The distance between two data points is calculated by Euclidean distance and is written as follows:

$$d_{ij} = \sqrt{\sum_{d=1}^{Dim} (x_i - x_j)^2} \tag{8}$$

where *Dim* represents the spatial dimension.

The local density ρ_i of the data point x_i is defined as follows:

$$\rho_i = \sum_j \chi(d_{ij} - d_c) \tag{9}$$

where d_c is the cutoff distance that needs to be set manually. $\chi(x) = \begin{cases} 1 & x \leq 0 \\ 0 & x > 0 \end{cases}$, indicates the number of data points with a distance less than d_c from the *i*-th data point as the true density of the data points x_i.

The distance δ_i from data point x_i to the nearest data point x_j whose local density is greater than x_i is defined as

$$\delta_i = \begin{cases} \min\left(d_{ij}\right), & \rho_j > \rho_i \\ \max\left(d_{ij}\right), & \rho_j \leq \rho_i \end{cases} \tag{10}$$

The value with the smallest distance x_i in the data points with larger local density than data point is found. If the local density of the data point x_i is the largest, the maximum distance between x_j and x_i is selected.

Combining the above parameters, the basic principles and steps of DPCA segmentation are described as follows:

Step 1: the Euclidean distance d_{ij} between any two data points are calculated;

Step 2: the cutoff distance d_c is set, and the local density ρ_i of data point x_i and distance δ_i are calculated;

Step 3: a clustering decision graph is drawn with the local density ρ_i as the horizontal axis and the distance δ_i as the vertical axis;

Step 4: the data points with high values of ρ_i and δ_i are marked as cluster centers for the decision graph and points with relatively small values ρ_i but relatively large values of δ_i as noise points;

Step 5: each remaining data point is assigned to its nearest neighbor and the class of the data point with greater density until all data points are allocated.

The algorithm idea of DPCA segmentation is much more intuitive. It can quickly find the density peak point and can efficiently complete the sample allocation and noise point recognition. The segmentation algorithm can quickly complete clustering without iteration and has high operating efficiency. However, the segmentation performance is restricted by the cutoff distance to a certain extent. The selection of cutoff distance mainly depends on the human experience. Therefore, the selection of the clustering center has certain subjective factors.

Kernel Density Estimation Optimization

The local density ρ of data points represents the distribution of neighboring data points in the area with the cutoff distance as the radius in the feature space. The selection of the cutoff distance will directly affect the clustering results. For the green target fruit, if the cutoff distance is not selected properly, it will be difficult to solve the boundary problem of the target image, so a more robust calculation method is needed. This study introduces a kernel density estimation optimization method to evaluate the local density of data points. This estimation is a non-parametric method, which can fully use its data characteristics to calculate, avoiding the influence of human prior knowledge. In other words, this method can get rid of the dependence on the cutoff distance d_c.

The local density of data points is defined by kernel density estimation:

$$\rho_i(x.y) = \frac{1}{nh} \sum_{i=1}^{n} K_h \left(\frac{x - x_i}{h}, \frac{y - y_i}{h} \right) \tag{11}$$

where K_h is the scaling kernel function, $h > 0$ is a smooth parameter. A Gaussian kernel function is adopted to smooth the peak function. Therefore, the scaling kernel function is

$$K\left(x, y\right) = \frac{1}{2\pi\sigma^2} \exp\left(-\frac{(x - x_i)^2 + \left(y - y_i\right)^2}{2\sigma^2} \right) \tag{12}$$

where, x_i, y_i are the mean values of the R-B area and G-B area of the *i*-th image block, respectively.

The result of using Gaussian kernel density to estimate the density distribution of the green fruit target image is shown in **Figure 9**.

Based on the local density ρ_i and distance δ_i, the cluster center is a data point with a larger local density and distance. Thus, a decision graph with the local density ρ_i as the horizontal axis and the distance δ_i as the vertical axis is constructed for finding cluster centers, as shown in **Figure 10**.

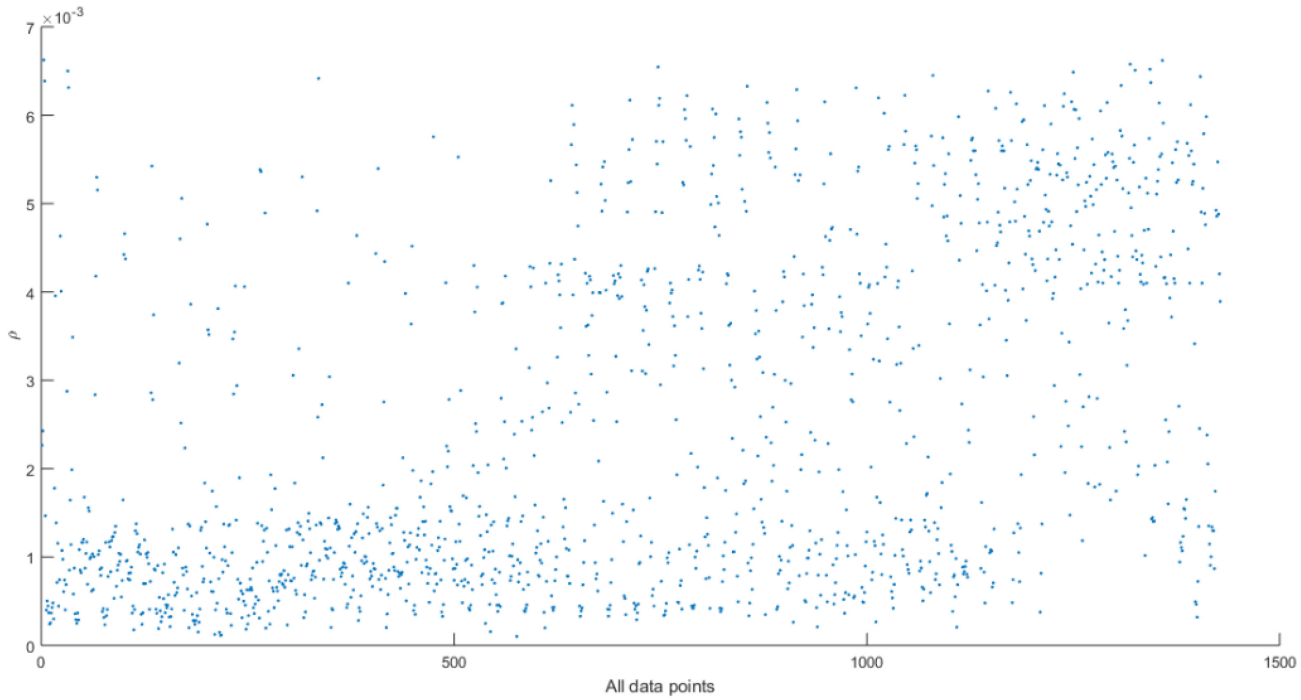


FIGURE 9 | Local density distribution of target image data points.

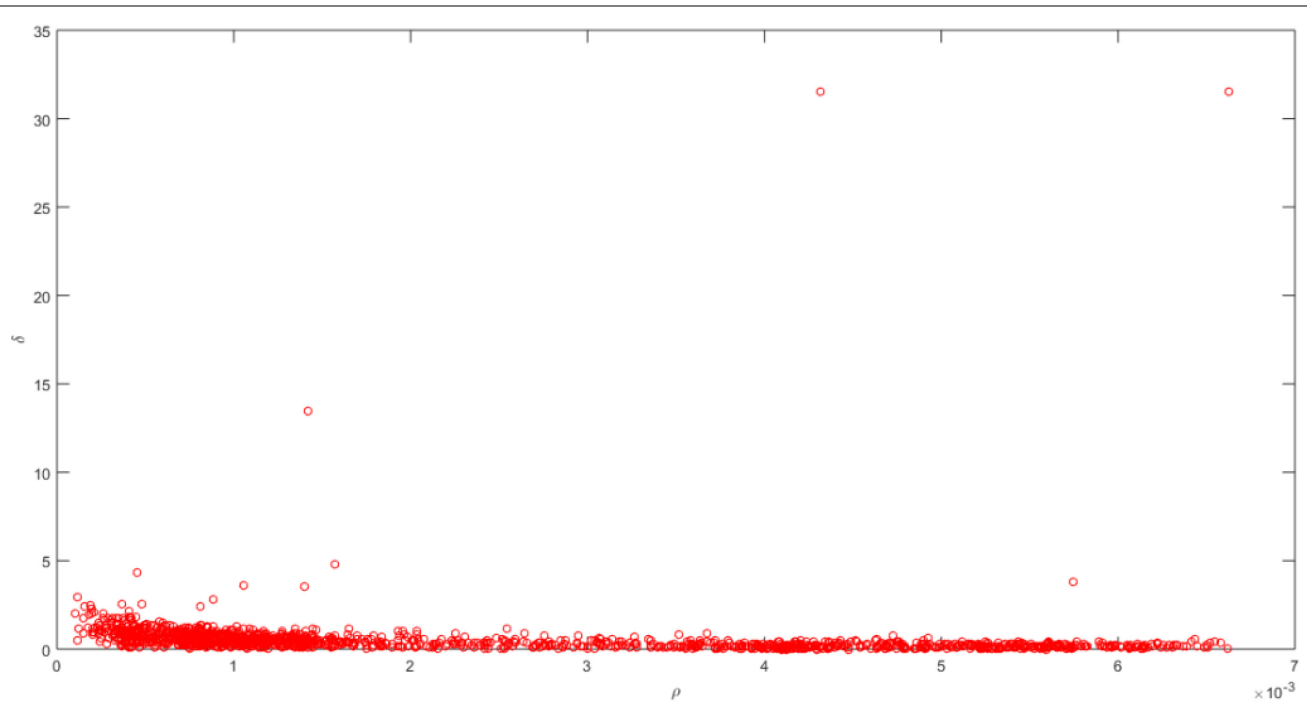


FIGURE 10 | Decision graph of target image clustering.

It can be seen in **Figure 10** that the data points with larger local density and distance in the decision graph are different cluster centers. However, the specific number of cluster centers needs to be artificially set according to actual problems. The data points close to the local density axis are cluster members and their distances are small. The data points close to the distant

axis are abnormal points, and the data points in the upper right corner are outliers.

Cluster Center Optimization

For the characteristics of the DPCA segmentation algorithm and the clustering process, the determination of non-cluster center data points does not completely depend on the cluster center, namely, the number of center points is not the number of final clusters. Considering that the complexity of the green apple image mainly lies in the complex structure between the target fruit and the background and the similarity of the two colors, it brings difficulties to the recognition of the target fruit. The segmentation of apple images is generally divided into three categories: target fruit, leaves, branches, and sky.

In this article, a double sorting algorithm is employed to automatically select the clustering center. Specifically, first, the local density value of all image blocks is sorted by descending and the first 20 image blocks are selected. Then, the distance values selected for 20 image blocks are sorted by descending order and the first 15 blocks are picked. Finally, the ratio of the number of pixels contained in each cluster and the number of pixels in the entire image is counted to analyze, as shown in **Figure 11**.

It can be seen from **Figure 11** that the first three categories account for 91.69% of the total number of pixels and the remaining categories of pixels account for a relatively low proportion. The clustering error of the target fruit area is brought by the boundary judgment and is related to the first 3 classes. The last 12 classes are regarded as the data non-allocation to the cluster center, which does not affect the target image segmentation. The segmentation effect obtained by the algorithm in this article is shown in **Figure 12**.

According to the segmentation result, the novel method completes the segmentation of the green target fruit and the background. The new method has a higher operating efficiency without iteration.

FITTING OF TARGET FRUIT

In the apple image, the target fruit can be regarded as a round-like shape, which guarantees the two elements of the recognizing and locating of the round-like target: accurate locating of the circle center and value of the radius. Although green apples have brought new challenges owing to the close color of the target fruit's skin and background, this study combines the depth and RGB image information of the green apple to solve the problem of recognizing and locating the green target fruit.

The center of the target fruit has been obtained in the depth image in section "The Center Location of the Target Fruit," and the segmented area of the target fruit is obtained in the RGB image in section "Target Area Segmentation." The above two results are merged. The target fruit area is the area where the center is located. In this study, the maximum value method is used to obtain the radius of the target fruit. Then, the contour of the target fruit is fitted to achieve the recognition and location of the target fruit. For a single target fruit, the maximum distance from the center of the circle to the edge of the region is directly selected. For overlapping target fruits, the minimum value from the extremum is used as the radius of the center. The process of

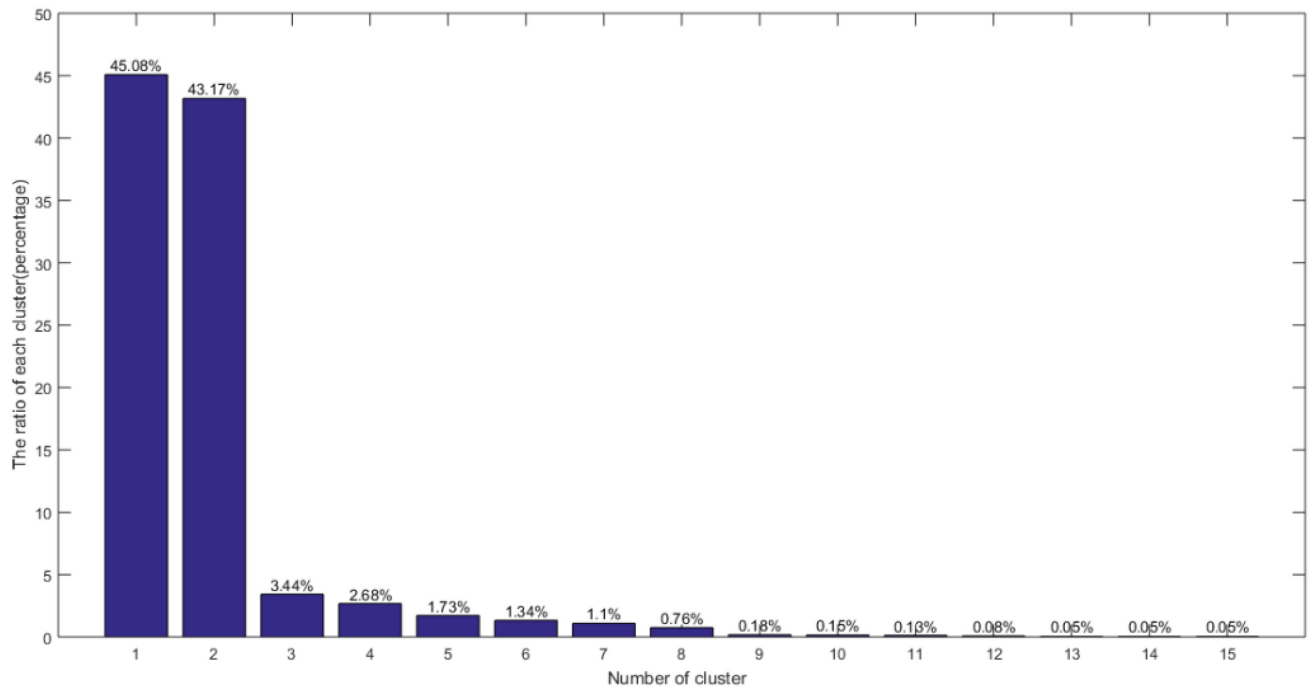


FIGURE 11 | The ratio of the number of pixels in each class to the total number of pixels.

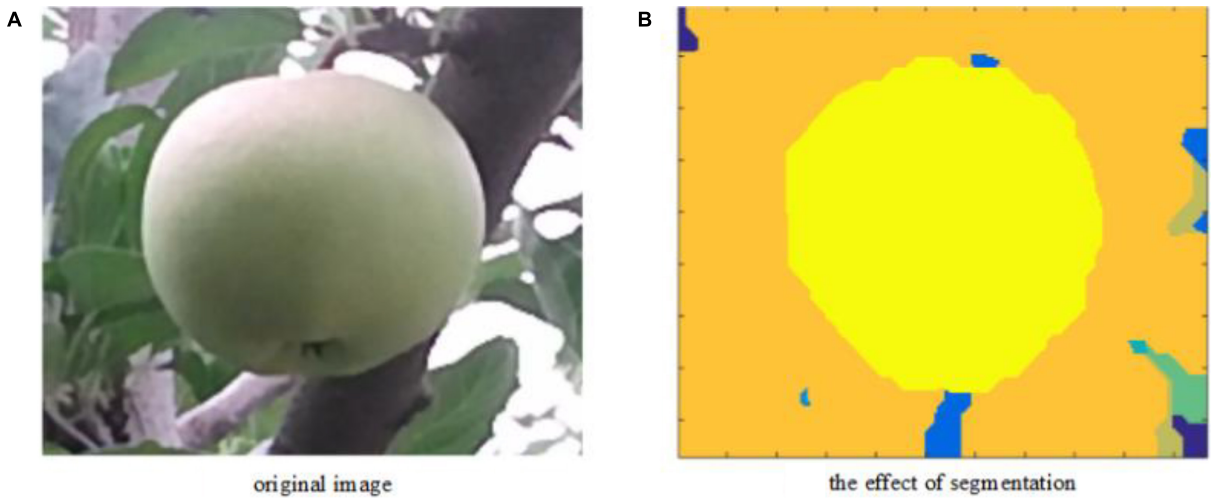


FIGURE 12 | Target image segmentation effect graph. **(A)** The left image represents the original RGB image and **(B)** the right image refers to the visualization result of the density peak clustering segmentation algorithm based on kernel density estimation corresponding to the original RGB image.

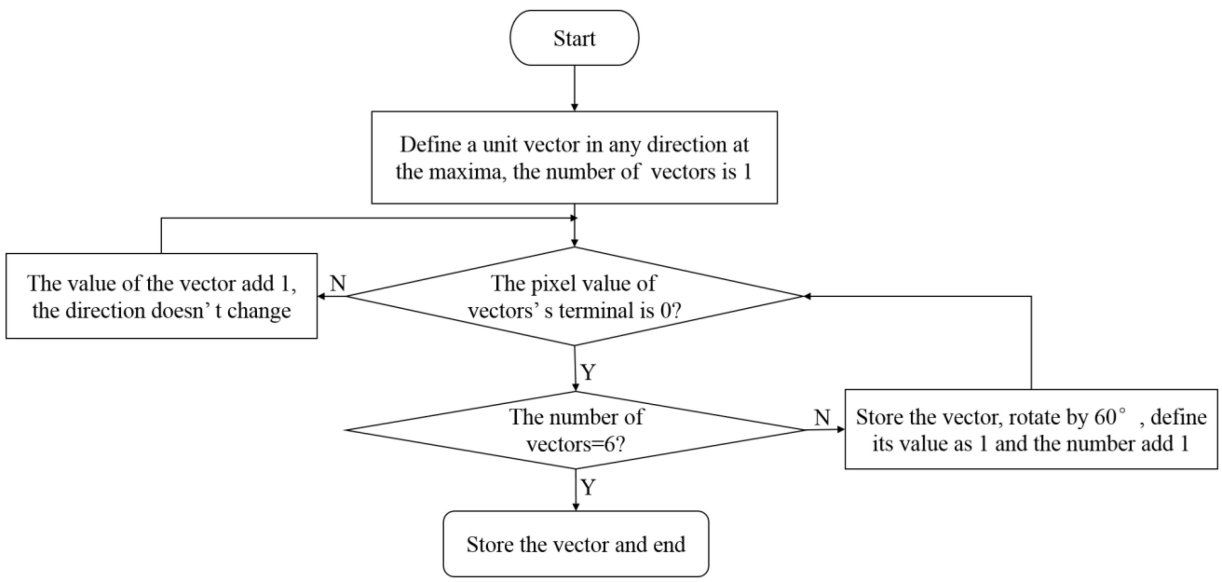


FIGURE 13 | The process of determining the radius of the overlapping target fruit.

determining the maximum radius of the overlapping target fruit is shown in **Figure 13**.

For overlapping target fruits, as shown in **Figure 13**, the segmented connected area containing the center of the circle is scanned to obtain six maximum vector lengths. The minimum value as the radius corresponding to the circle center is applied to realize the recognition and location.

Therefore, the center and radius of the circle have been determined, and the contours of the target fruit are fitted to complete the recognition and location of the target fruit, as shown in **Figure 14**.

After the new method completes the segmentation, there is no need to extract the features of the target fruit segmentation area, and there is no need to design a classifier. The target fruit can be determined by the central location. It can be concluded that the recognition accuracy and efficiency of the novel method are relatively high.

RESULTS AND DISCUSSION

Based on the above theoretical analysis, this research proposes a new method that can achieve rapid and accurate recognition and location of the target fruit of green apples. To better verify the effectiveness of the new algorithm, the following experiments are conducted.

FIGURE 14 | Fitting effect of target fruit.

Result Assessment

The collected 300 pairs of green apple images were used in the experiment, including 526 target fruits. A single target fruit, occlusions of branches and leaves, and overlapping fruits are 287, 89, and 150, respectively. Recognition rate and running time are used to evaluate the recognition and location method. The recognition rates are calculated, including overall recognition rate, single target fruit recognition rate, obscured target fruit recognition rate, and overlapping target fruit recognition rate. The results of our optimized method with different growth postures are presented in **Table 1**. Note that the experimental operating platform of this study is as follows: the host configuration is equipped with an Intel Core i5-5257 CPU clocked at 2.7GHz and 4GB of memory; the operating environment is MATLAB R2015a installed with a 64-bit Windows 10 operating system. From **Table 1**, the total recognition rate of our method can reach 96.96%, where the recognition of fruits without obscurity is 99.3%. It can be concluded that our method has a high recognition rate.

To better verify the performance of the new method, the experiment is compared with the method in the literature (Li et al., 2019; Lv et al., 2019a; Jiao et al., 2020). Meanwhile, the performance of the new method is evaluated from the above indicators. The results of comparative experiments with the literature (Li et al., 2019; Lv et al., 2019a; Jiao et al., 2020) are listed in **Table 1**.

TABLE 1 | Correct recognition rate of three type's fruits (%).

Model	Fruit without obscured	Obscured fruit	Overlapping fruit	Total recognition rate
Ours	**99.30**	**95.33**	**92.13**	**96.96**
Ref. 16	97.56	91.33	86.51	93.92
Ref. 12	97.91	89.33	85.39	93.35
Ref. 36	96.17	90.67	89.89	93.53

Note that the bold value represents the optimal value under a specific evaluation metric.

TABLE 2 | Comparison of the recognition performance of each algorithm.

Model	Recognition time (ms)	Total recognition rate (%)
Ours	**897**	**96.96**
Ref. 16	1269	93.92
Ref. 12	1505	93.35
Ref. 36	1627	93.53

Note that the bold value represents the optimal value under a specific evaluation metric.

Further, the operating efficiency and recognition accuracy of the whole new algorithm are evaluated, and the recognition time and recognition accuracy are calculated. The results are listed in **Table 2**.

It can be found from **Tables 1**, **2** that the efficiency of recognition and location of a single unobstructed fruit with a single branch and leaf obscured are the highest, and the recognition rate for overlapping target fruits is slightly lower. Based on the overall comparison effect, the recognition rate and operation efficiency have been greatly improved. The new method in this article is better than other methods.

Visualization Result of Location and Recognition

Due to the variable natural growth posture of apple fruit, the shooting angle, and other factors, the collected target fruit presents three postures: single unobstructed, obscured by branches and leaves, and overlapping fruits. According to the new method in this article, three types of samples are tested, and the recognition results obtained are shown in **Figure 15**. We can find that a single unobstructed fruit, overlapping fruit, and fruit occluded by branches and leaves can all recognize and locate the target fruit.

Result Analysis

It can be seen in **Figure 15** and **Tables 1**, **2** that for the problem of green target fruit recognition and location, the new method has greatly improved the accuracy and operation efficiency of the recognition and location. The recognition rate of a single target fruit can reach almost 100%, indicating that the novel method is feasible.

In general, the novel method is designed without iteration during the clustering and segmentation process, without features of the segmentation area and a classifier during the recognition process. Therefore, the performance of the novel method is greatly improved. In terms of recognition and location accuracy, the new method can accurately locate the center and radius of the circle. The performance is also significantly improved.

Limitation Discussion

Although the overall recognition rate of the target fruit is relatively high, there is still a small problem: the recognition accuracy of excessively overlapping target fruits is low owing

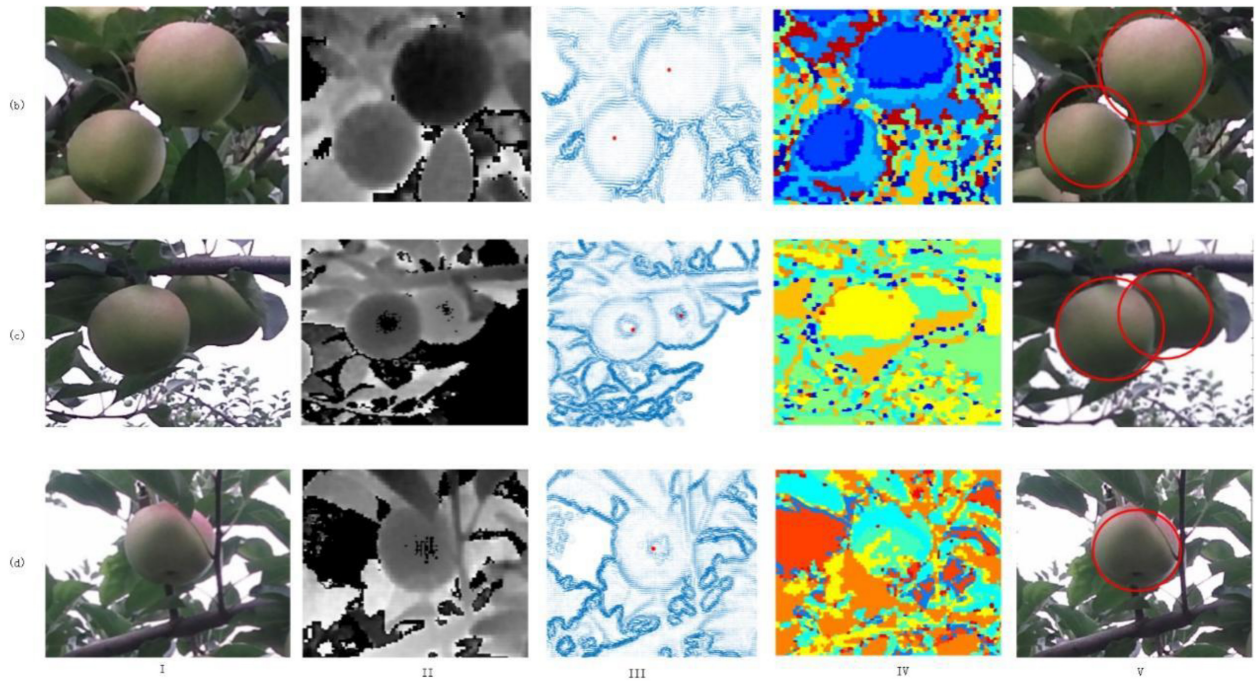


FIGURE 15 | Three types of target fruit recognition and location effect (I) RGB space image, (II) depth image, (III) center position effect image, (IV) segmentation rendering, (V) identify and location.

to the occlusion of branches and leaves for overlapping target fruits. Specifically, when the maximum value is used to find the radius corresponding to the circle center in the connected area of the segmented overlapping target fruit, the minimum value is scanned for the occluded area of branches and leaves, which makes it difficult to obtain the optimal radius. An excessively overlapping apple sample is shown in **Figure 16**.

For the problem of difficulty in locating the center of excessively overlapping fruit, we will consider optimizing the method of locating the center to reduce the dependence on depth information during the positioning process, that is, using a small amount of gradient information to achieve the method of locating the center of the circle. In addition, the RGB image segmentation method will be optimized to obtain smoother segmentation boundaries.

FIGURE 16 | The visualization results of an excessively overlapping apple.

CONCLUSION

Focusing on the problem of recognition and location of green apples, this article proposes a fast and accurate location and recognition method based on the fusion of depth images and RGB images. First, the gradient field information in the depth image is employed to draw an iso-level contour map and gradient vectors of the target image. All gradient vectors are rotated in the same direction to obtain different vorticity. The three-dimensional depth image is projected to the two-dimensional plane to find the center of vorticity. The center of vorticity is the center of the target fruit, achieving an accurate location of the center of the target fruit. Second, a density peak clustering algorithm is used to segment the target fruit in the RGB image. To get rid of the subjective factors of selecting a cutoff distance, the Gaussian kernel density estimation is applied to optimize the algorithm. For the cluster centers, a double sort method is utilized to select the cluster centers automatically to achieve efficient segmentation of target images. Finally, the results of the two steps are merged and the segmentation area of the circle center is just the target fruit area. The maximum value method is applied to obtain the

target fruit radius, and the contour of the target fruit is fitted to complete the efficient and accurate recognition and location of the target fruit.

The proposed method is presented without iteration during the process of center location and clustering segmentation, without features of the segmentation area and a classifier in the recognition process. Therefore, the running efficiency of the new method is significantly improved. In the process of center location and radius calculation, the error is relatively small, and thus, the recognition and location accuracy of the new method is higher. Experimental results also show that the new method has greatly improved the accuracy and operating efficiency of recognition and location. In addition, our method is a lightweight method that does not require high-performance servers for computing, which can easily to transplant and embedded into the hardware environment of the fruit picking robot or applied to the automatic monitoring process of fruit growth. Therefore, the new method can solve the problem of recognizing and locating green target fruit well. The proposed method can be further extended to the problem of rapid and accurate recognition and location of spherical fruits, which can be used in the field of machine harvesting or yield estimation of spherical fruits. For the task of fruit recognition with overlapping targets occluded by branches and leaves, the radius calculation of the new method is susceptible to be interference, which will be the focus of future research.

DATA AVAILABILITY STATEMENT

The original contributions presented in this study are included in the article/supplementary material, further inquiries can be directed to the corresponding authors.

AUTHOR CONTRIBUTIONS

MS: conceptualization and writing—original draft preparation. LX: data curation, software, and validation. RL: funding acquisition and visualization. YL: methodology, software, and validation. WJ: conceptualization, funding acquisition, and writing—reviewing and editing. All authors contributed to the article and approved the submitted version.

FUNDING

This work was supported by the National Nature Science Foundation of China (no. 21978139), the Natural Science Foundation of Shandong Province in China (ZR2020MF076 and ZR2019MB030), the Focus on Research and Development Plan in Shandong Province (no. 2019GNC106115), and the Taishan Scholar Program of Shandong Province of China.

REFERENCES

Achanta, R., Shaji, A., Smith, K., Lucchi, A., Fua, P., and Süsstrunk, S. (2012). SLIC superpixels compared to state-of-the-art superpixel methods. *IEEE Trans. Pattern Anal. Mach. Intell.* 34, 2274–2282. doi: 10.1109/TPAMI.2012.120

Bargoti, S., and Underwood, J. P. (2017). Image segmentation for fruit detection and yield estimation in apple orchards. *J. Field Robot.* 34, 1039–1060. doi: 10.1002/rob.21699

Barnea, E., Mairon, R., and Ben-Shahar, O. (2016). Colour-agnostic shape-based 3D fruit detection for crop harvesting robots. *Biosyst. Eng.* 146, 57–70. doi: 10.1016/j.biosystemseng.2016.01.013

Behera, S. K., Rath, A. K., Mahapatra, A., and Sethy, P. K. (2020). Identification, classification & grading of fruits using machine learning & computer intelligence: a review. *J. Ambient Intell. Humaniz. Comput.* 11, 1–11. doi: 10.1007/s12652-020-01865-8

Biffi, L. J., Mitishita, E., Liesenberg, V., Santos, A. A. d, Gonçalves, D. N., Estrabis, N. V., et al. (2020). Atss deep learning-based approach to detect apple fruits. *Remote Sensing* 13:54. doi: 10.3390/rs13010054

Chen, Z., Wu, R., Lin, Y., Li, C., Chen, S., Yuan, Z., et al. (2022). Plant disease recognition model based on improved YOLOv5. *Agronomy* 12:365. doi: 10.3390/agronomy12020365

Choi, D., Lee, W. S., Schueller, J. K., Ehsani, R., Roka, F., and Diamond, J. (2017). "A performance comparison of RGB, NIR, and depth images in immature citrus detection using deep learning algorithms for yield prediction," in *Proceedings of the 2017 ASABE Annual International Meeting*, (St. Joseph: American Society of Agricultural and Biological Engineers). doi: 10.13031/aim.201700076

Gené-Mola, J., Vilaplana, V., Rosell-Polo, J. R., Morros, J.-R., Ruiz-Hidalgo, J., and Gregorio, E. (2019). Multi-modal deep learning for Fuji apple detection using RGB-D cameras and their radiometric capabilities. *Comput. Electron. Agric.* 162, 689–698. doi: 10.1016/j.dib.2019.104289

Genno, H., and Kobayashi, K. (2019). Apple growth evaluated automatically with high-definition field monitoring images. *Comput. Electron. Agric.* 164, 104895. doi: 10.1016/j.compag.2019.104895

Häni, N., Roy, P., and Isler, V. (2020a). A comparative study of fruit detection and counting methods for yield mapping in apple orchards. *J. Field Robot.* 37, 263–282. doi: 10.1002/rob.21902

Häni, N., Roy, P., and Isler, V. (2020b). Minneapple: a benchmark dataset for apple detection and segmentation. *IEEE Robot. Autom. Lett.* 5, 852–858. doi: 10.1109/LRA.2020.2965061

He, L., Fu, H., Karkee, M., and Zhang, Q. (2017). Effect of fruit location on apple detachment with mechanical shaking. *Biosyst. Eng.* 157, 63–71. doi: 10.1016/j.biosystemseng.2017.02.009

Ji, W., Gao, X., Xu, B., Chen, G., and Zhao, D. (2020). Target recognition method of green pepper harvesting robot based on manifold ranking. *Comput. Electron. Agric.* 177:105663.

Jia, W., Tian, Y., Luo, R., Zhang, Z., Lian, J., and Zheng, Y. (2020b). Detection and segmentation of overlapped fruits based on optimized mask R-CNN application in apple harvesting robot. *Comput. Electron. Agric.* 172:105380. doi: 10.1016/j.compag.2020.105380

Jia, W., Zhang, Y., Lian, J., Zheng, Y., Zhao, D., and Li, C. (2020a). Apple harvesting robot under information technology: a review. *Int. J. Adv. Robot. Syst.* 17:1729881420925310. doi: 10.1177/1729881420925310

Jiao, Y., Luo, R., Li, Q., Deng, X., Yin, X., Ruan, C., et al. (2020). Detection and localization of overlapped fruits application in an apple harvesting robot. *Electronics* 9:1023. doi: 10.3390/electronics9061023

Kamilaris, A., and Prenafeta-Boldú, F. X. (2018). Deep learning in agriculture: a survey. *Comput. Electron. Agric.* 147, 70–90. doi: 10.1016/j.compag.2018.02.016

Kang, H., and Chen, C. (2020). Fruit detection, segmentation and 3D visualisation of environments in apple orchards. *Comput. Electron. Agric.* 171:105302. doi: 10.1016/j.compag.2020.105302

Koirala, A., Walsh, K. B., Wang, Z., and McCarthy, C. (2019). Deep learning–Method overview and review of use for fruit detection and yield estimation. *Comput. Electron. Agric.* 162, 219–234. doi: 10.1016/j.compag.2019.04.017

Li, B., Long, Y., and Song, H. (2018). Detection of green apples in natural scenes based on saliency theory and Gaussian curve fitting. *Int. J. Agric. Biol. Eng.* 11, 192–198. doi: 10.25165/j.ijabe.20181101.2899

Li, D.-h, Hui, Z., and Xiao, Y. (2019). Overlapping green apple recognition based on improved spectral clustering. *Spectrosc. Spectr. Anal.* 39:2974.

Li, Q., Jia, W., Sun, M., Hou, S., and Zheng, Y. (2021). A novel green apple segmentation algorithm based on ensemble U-Net under complex orchard environment. *Comput. Electron. Agric.* 180:105900. doi: 10.1016/j.compag.2020.105900

Liu, G., Nouaze, J. C., Touko Mbouembe, P. L., and Kim, J. H. (2020). YOLO-tomato: a robust algorithm for tomato detection based on YOLOv3. *Sensors* 20:2145. doi: 10.3390/s20072145

Lv, J., Wang, F., Xu, L., Ma, Z., and Yang, B. (2019a). A segmentation method of bagged green apple image. *Sci. Hortic.* 246, 411–417. doi: 10.1016/j.scienta.2018.11.030

Lv, J., Wang, Y., Ni, H., Wang, Q., Rong, H., Ma, Z., et al. (2019b). Method for discriminating of the shape of overlapped apple fruit images. *Biosyst. Eng.* 186, 118–129.

Maheswari, P., Raja, P., Apolo-Apolo, O. E., and Pérez-Ruiz, M. (2021). Intelligent fruit yield estimation for orchards using deep learning based semantic segmentation techniques—a review. *Front. Plant Sci.* 12:684328. doi: 10.3389/fpls.2021.684328

Rodriguez, A., and Laio, A. (2014). Clustering by fast search and find of density peaks. *Science* 344, 1492–1496. doi: 10.1126/science.1242072

Si, Y., Liu, G., and Feng, J. (2015). Location of apples in trees using stereoscopic vision. *Comput. Electron. Agric.* 112, 68–74. doi: 10.1016/j.compag.2015.01.010

Silwal, A., Davidson, J. R., Karkee, M., Mo, C., Zhang, Q., and Lewis, K. (2017). Design, integration, and field evaluation of a robotic apple harvester. *J. Field Robot.* 34, 1140–1159.

Sun, S., Jiang, M., He, D., Long, Y., and Song, H. (2019). Recognition of green apples in an orchard environment by combining the GrabCut model and Ncut algorithm. *Biosyst. Eng.* 187, 201–213. doi: 10.1016/j.biosystemseng.2019.09.006

Tang, Y., Chen, M., Wang, C., Luo, L., Li, J., Lian, G., et al. (2020). Recognition and localization methods for vision-based fruit picking robots: a review. *Front. Plant Sci.* 11:510. doi: 10.3389/fpls.2020.00510

Wang, D., He, D., Song, H., Liu, C., and Xiong, H. (2019). Combining SUN-based visual attention model and saliency contour detection algorithm for apple image segmentation. *Multimed. Tools Appl.* 78, 17391–17411. doi: 10.1007/s11042-018-7106-y

Wu, F., Duan, J., Chen, S., Ye, Y., Ai, P., and Yang, Z. (2021). Multi-target recognition of bananas and automatic positioning for the inflorescence axis cutting point. *Front. Plant Sci.* 12:705021. doi: 10.3389/fpls.2021.705021

Zhang, C., Zhang, J., Zhang, J., and Li, W. (2014). Recognition of green apple in similar background. *Nongye Jixie Xuebao Trans. Chin. Soc. Agric. Machinery* 45, 277–281.

Zhang, J., Yang, B., Geng, N., and Huang, L. (2017). An obstacle detection system based on monocular vision for apple orchardrobot. *Int. J. Robot. Autom.* 32, 639–648. doi: 10.2316/Journal.206.2017.6.206-5036

Zhao, Z.-Q., Zheng, P., S.-t, Xu, and Wu, X. (2019). Object detection with deep learning: a review. *IEEE Trans. Neural Netw. Learn. Syst.* 30, 3212–3232. doi: 10.1109/TNNLS.2018.2876865

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

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

frontiers | Frontiers in Plant Science

ORIGINAL RESEARCH
published: 20 June 2022
doi: 10.3389/fpls.2022.897739

Intelligent Monitoring System of Migratory Pests Based on Searchlight Trap and Machine Vision

Guojia Sun[1], Shuhua Liu[2], Haolun Luo[1], Zelin Feng[1], Baojun Yang[2], Ju Luo[2], Jian Tang[2], Qing Yao[1]* and Jiajun Xu[1]*

[1]School of Information Science and Technology, Zhejiang Sci-Tech University, Hangzhou, China, [2]State Key Laboratory of Rice Biology, China National Rice Research Institute, Hangzhou, China

OPEN ACCESS

Edited by:
Lei Shu,
Nanjing Agricultural University, China

Reviewed by:
Sijia Yu,
Rutgers, The State University of New Jersey–Busch Campus, United States
Tianyu Xi,
Institute of Zoology (CAS), China

***Correspondence:**
Shuhua Liu
liushuhua@caas.cn
Qing Yao
q-yao@126.com

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

Received: 16 March 2022
Accepted: 24 May 2022
Published: 20 June 2022

Citation:
Sun G, Liu S, Luo H, Feng Z, Yang B, Luo J, Tang J, Yao Q and Xu J (2022) Intelligent Monitoring System of Migratory Pests Based on Searchlight Trap and Machine Vision. Front. Plant Sci. 13:897739. doi: 10.3389/fpls.2022.897739

Three species of rice migratory pests (*Cnaphalocrocis medinalis, Sogatella furcifera*, and *Nilaparvata lugens*) cause severe yield and economic losses to rice food every year. It is important that these pests are timely and accurately monitored for controlling them and ensuring food security. Insect radar is effective monitoring equipment for migratory pests flying at high altitude. But insect radar is costly and has not been widely used in fields. Searchlight trap is an economical device, which uses light to trap migratory pests at high altitude. But the trapped pests need to be manually identified and counted from a large number of non-target insects, which is inefficient and labor-intensive. In order to replace manual identification of migratory pests, we develop an intelligent monitoring system of migratory pests based on searchlight trap and machine vision. This system includes a searchlight trap based on machine vision, an automatic identification model of migratory pests, a Web client, and a cloud server. The searchlight trap attracts the high-altitude migratory insects through lights at night and kills them with the infrared heater. All trapped insects are dispersed through a multiple layers of insect conveyor belts and a revolving brush. The machine vision module collects the dispersed insect images and sends them to the cloud server through 4G network. The improved model YOLO-MPNet based on YOLOv4 and SENet channel attention mechanism is proposed to detect three species of migratory pests in the images. The results show that the model effectively improves the detection effect of three migratory pests. The precision is 94.14% for *C. medinalis*, 85.82% for *S. furcifera*, and 88.79% for *N. lugens*. The recall is 91.99% for *C. medinalis*, 82.47% for *S. furcifera*, and 85.00% for *N. lugens*. Compared with some state-of-the-art models (Faster R-CNN, YOLOv3, and YOLOv5), our model shows a low false detection and missing detection rates. The intelligent monitoring system can real-timely and automatically monitor three migratory pests instead of manually pest identification and count, which can reduce the technician workload. The trapped pest images and historical data can be visualized and traced, which provides reliable evidence for forecasting and controlling migratory pests.

Keywords: searchlight trap, rice migratory pests, intelligent monitoring, deep learning, machine vision

INTRODUCTION

According to the Food and Agriculture Organization of the United Nations, the annual potential loss of crop yield caused by pests is about 30% worldwide. Migratory pests are among the most harmful, as they can cause great disasters in a short period of time (Hu et al., 2020). Considering the long-distance migratory ability of migratory pests, dynamic monitoring of migratory pests' population is crucial for timely and effective pest management. Dynamic monitoring of pest populations includes adult monitoring, field pest egg survey, and damage symptom investigation. Among them, timely monitoring of adult occurrence time and quantity is the basis of effective pest management (Jiang et al., 2021). At present, the adult monitoring equipment of migratory pests mainly includes insect radar and light trap. Insect radar mainly indirectly monitors pest species and quantity through calculating the insect flapping wing frequency, the body shape, and size of each insect in radar images (Feng, 2011; Zhang et al., 2017). In fact, the insect wing flapping frequency is related to insect instar and flight environment temperature. Insects with the same shape and size may be different insect species. Consequently, it is difficult to accurately identify the insect species which becomes a major obstacle to the widespread application of insect radar in fields for migratory pest forecasting (Feng, 2003). As an important tool for monitoring agricultural pests, light traps can be divided into two types (Yang et al., 2017). One is for trapping pests in fields, named ground light trap. The other is for trapping pests in high-altitude, named searchlight trap. Compared with the ground light trap, searchlight trap shows superiority in monitoring migratory pests, such as larger biomass, longer monitoring period, and more obvious fluctuation curve of pest quantity (Jiao et al., 2017; Shang, 2017; Qin, 2019). From 2014, searchlight traps (using metal halide lamps, bulb light source wavelength of 500–600 nm, and power of 1,000 W) have been used to monitor regional migratory pests and obtained good monitoring results (Jiang et al., 2016). However, the identification and count of pests trapped by the searchlight traps still needs to be carried out manually. This manual method requires high professional skills and spends much time, which causes low efficiency, high labor intensity and non-timely data application (Song et al., 2021; Yan et al., 2021).

With the development of machine vision, there has been some progress in pest detection and recognition studies based on images. Qiu et al. (2007) used the automatic threshold segmentation, feature extraction, and BP neural network classifier method to identify nine species of field pests. Based on the morphology and color features of pests, Han and He (2013) developed a support vector machine classifier to automatically identify six species of field pests. Zou (2013) adopted four different methods to extract shape features of rice planthoppers to improve the accuracy of pest identification. Yao et al. (2021a) proposed an automatic pest detection method based on improved CornerNet, which effectively improved the detection effect of rice planthoppers on light-trap insect images. Feng (2020) proposed YOLO-pest model to detect three species of *Cnaphalocrocis medinalis*, *Chilo suppressalis*, and *Sesamia inferens*, which reduced false detection and missing detection caused by insect adhesions in images. In order to improve the detection precision of light-trap insects, Yao et al. (2021b) proposed a bilinear attention network to identify similar light-tap pests. But there are no reports about intelligent searchlight traps based on machine vision and its pest identification methods. The challenges of pest identification from searchlight traps are (1) to timely disperse insects for collecting high-quality pest images, (2) to accurately identify the small size of pests, (3) to accurately distinguish those similar pests, and (4) to identify target pests from a large number of non-target insects.

To automatically identify and count rice migratory pests trapped by searchlight traps, we design an intelligent monitoring system of migratory pests based on searchlight trap and machine vision. The system can realize the automatic identification and count of three species of rice migratory pests (*C. medinalis, Sogatella furcifera*, and *Nilaparvata lugens*) attracted by searchlight trap.

MATERIALS AND METHODS

Intelligent Monitoring System of Migratory Pests

The Intelligent monitoring system of migratory pests consists of an intelligent searchlight trap based on machine vision, an automatic identification model of migratory pests, a Web client, and a cloud server. The searchlight trap firstly attracts and kills insects, then the machine vision module disperses insects and captures images. After these images are uploaded to the server, the server runs the model to identify migratory pests in the images. Finally, the identification results of pests are presented to the Web client. **Figure 1** shows the system construction.

The searchlight trap is mainly composed of a searchlight, an infrared heater module, and an insect collection box. When the equipment works at night, those flying insects within a high altitude of 500 m can be attracted. After the insects drop into the equipment from the top of the searchlight trap, they are killed and dried by the infrared heater module.

The machine vision module includes multilayer insect conveyor belts, Android PAD, industrial camera (MV-CE200-10GC), area light source, and network transfer module. Firstly, the dead insects fall into a vibrating slope controlled by a vibration motor for dispersing insects. Then, insects are dispersed onto the first layer of conveyor belts for further dispersion. Before falling onto the second layer of belt, the big insects are left and small insects are dispersed to the third layer of belt by a revolving brush for avoiding big insects covering small insects. When the big insects are transmitted onto the second layer, the camera takes photos of insects on the third layer of belt. The images are uploaded to the cloud server through the network transfer module on the Android PAD. The Android PAD is equipped with a special program to display images and photograph information (photograph time, image number, etc.) in real-time. The parameters can be manually set on the screen to control the photograph. Finally, all insects fall into the insect collection box at the bottom after they are photographed.

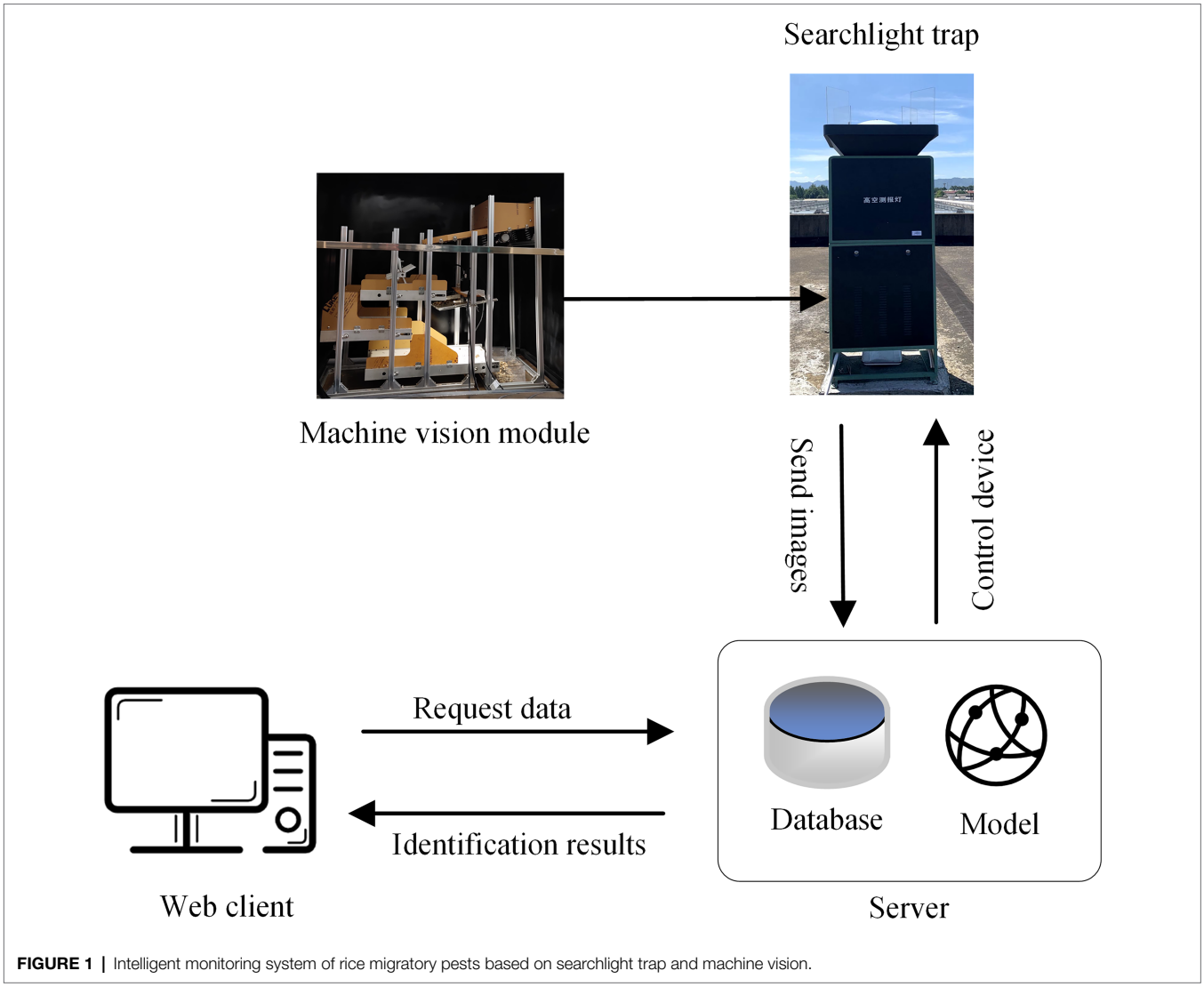


FIGURE 1 | Intelligent monitoring system of rice migratory pests based on searchlight trap and machine vision.

TABLE 1 | The number of migratory pests on images.

Datasets	Image number	Pest number			
		C. medinalis	*S. furcifera*	*N. lugens*	Interference pests
Training sets	2,187	73,146	90,126	59,250	8,487
Test sets	243	6,993	9,006	5,850	822

The Image Dataset

The intelligent searchlight trap was installed in the paddy fields in Fuyang District, Zhejiang Province. 2,430 images with rice migratory pests were collected in 2021. The size of an image is 5,472 × 3,648 pixels.

The migratory pest images were divided into a training set and a testing set in the ratio of 9:1. We used the LabelImg tool to annotate three species of migratory pests (*C. medinalis*, *S. furcifera*, and *N. lugens*) in images. The classification information and coordinate information of the labeled region were saved in the corresponding XML file. The searchlight trap caught many non-target insects as well. Some of them are very similar to the target pests visually, which leads to false detection. These non-target insects are called interference pests in this paper. The information of dataset is shown in **Table 1**.

Image Preprocessing

Image Data Enhancement

As we all know, the larger the dataset, the better the generalization performance for deep learning methods. To improve the robustness and generalization ability of the automatic identification model of migratory pests, we use some image processing methods to increase the number of images for training models. These methods include image left and right mirror, 90° rotating image, image equalization, and adding Gaussian noise (Lee, 1980). The algorithm functions of these methods are called from OpenCV library. Finally, the training sample number is increased by four times. The data enhancement results are shown in **Figure 2**.

Overlapping Sliding Window Method

Among three migratory pests, the size of two planthopper pests is about 3–5 mm, accounting for about 0.06% of the original image size. Due to the small area proportion of one pest in one image, the feature extraction network cannot extract effective features, which results in a concerning problem of missing detection of planthoppers. We adopt the overlapping sliding window processing method (OSW; Yao et al., 2021a) to improve the area proportion of each target pest in sub-images. The method can reduce missing detection and improve detection precision.

The original image is 5,472 × 3,648 pixels. We quarter the size of the original image and then add the length of the circumscribed rectangle of the largest pest (300 pixels) in the image to determine the size of the fixed window as 1,668 × 1,212 pixels. During detection, the sliding window slides towards the center from all sides. The order of movement is to move from the position (1) slides to (2) and (3) respectively, then slides from (3) to position (4). **Figure 3** shows the implementation of overlapping sliding window processing method. It cuts out the image in the window to become a new subimage when sliding. The size of the new subimage is smaller than the original image, but the size of the pests in the subimage has not changed, so the area ratio of each target pest has increased. The change from small target into "large" target contributes to extract the features of the target more efficiently. In the example figure, the sliding window takes the image at position (1) and then the picture at position (2). An overlapping area of pest length is left between the two subimages, which ensures that each pest will be fully learned and detected at least once. In this way, the missing detection is reduced and the number of data sets can be increased without destroying the integrality of the insect body, which is beneficial for improving the detection precision of the model. If a target pest happens to appear on the boundary of the sliding window, part of the pest body appears in the sliding window and it may be detected by the detection box. This problem is subsequently solved by the target detection box suppression method.

Detection Model of Rice Migratory Pests

Model Network Framework

Typical single-stage object detection models include the YOLO series (Redmon et al., 2016; Redmon and Farhadi, 2017, 2018), SSD (Liu et al., 2016), etc. The YOLO model is known for both speed and precision.

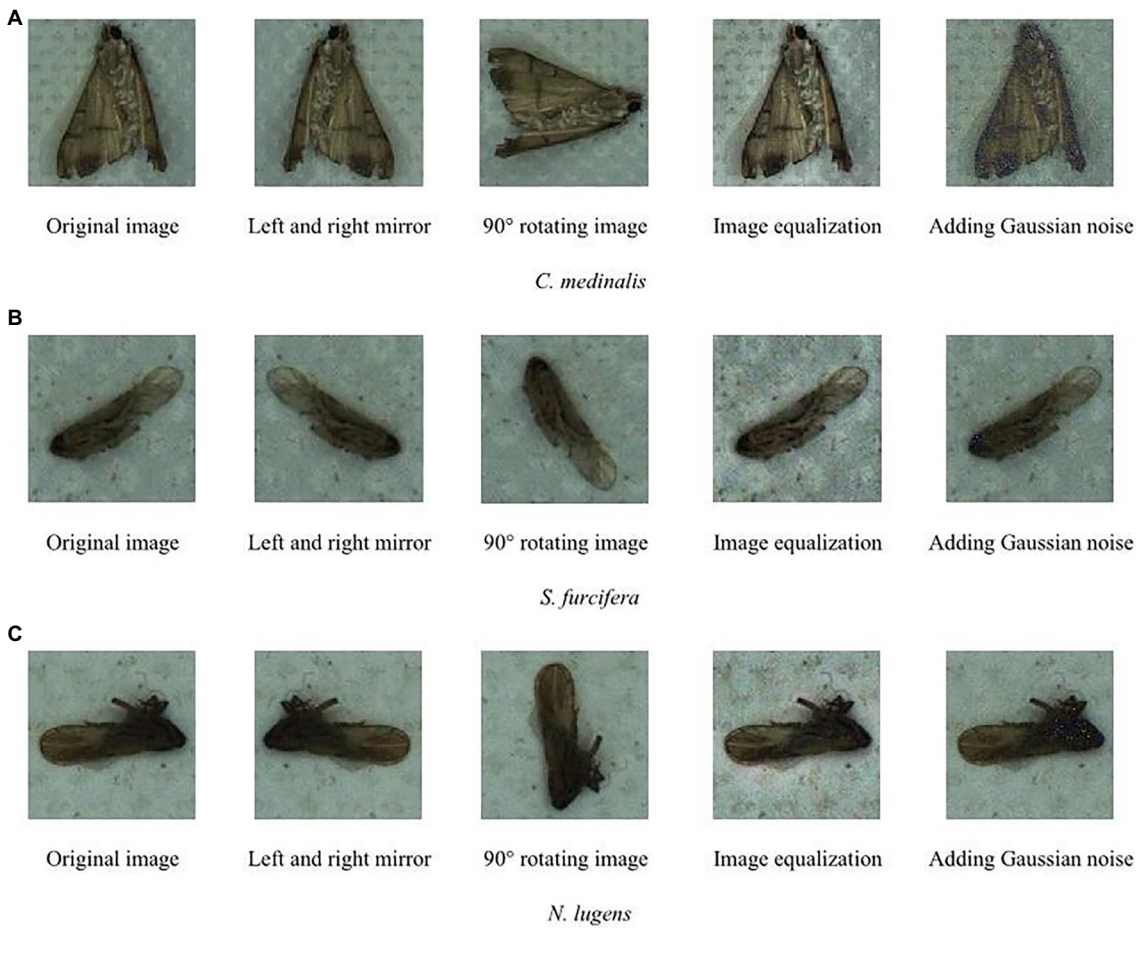


FIGURE 2 | Data enhancement of target pest images **(A)** *C. medinalis*, **(B)** *S. furcifera*, and **(C)** *N. lugens*.

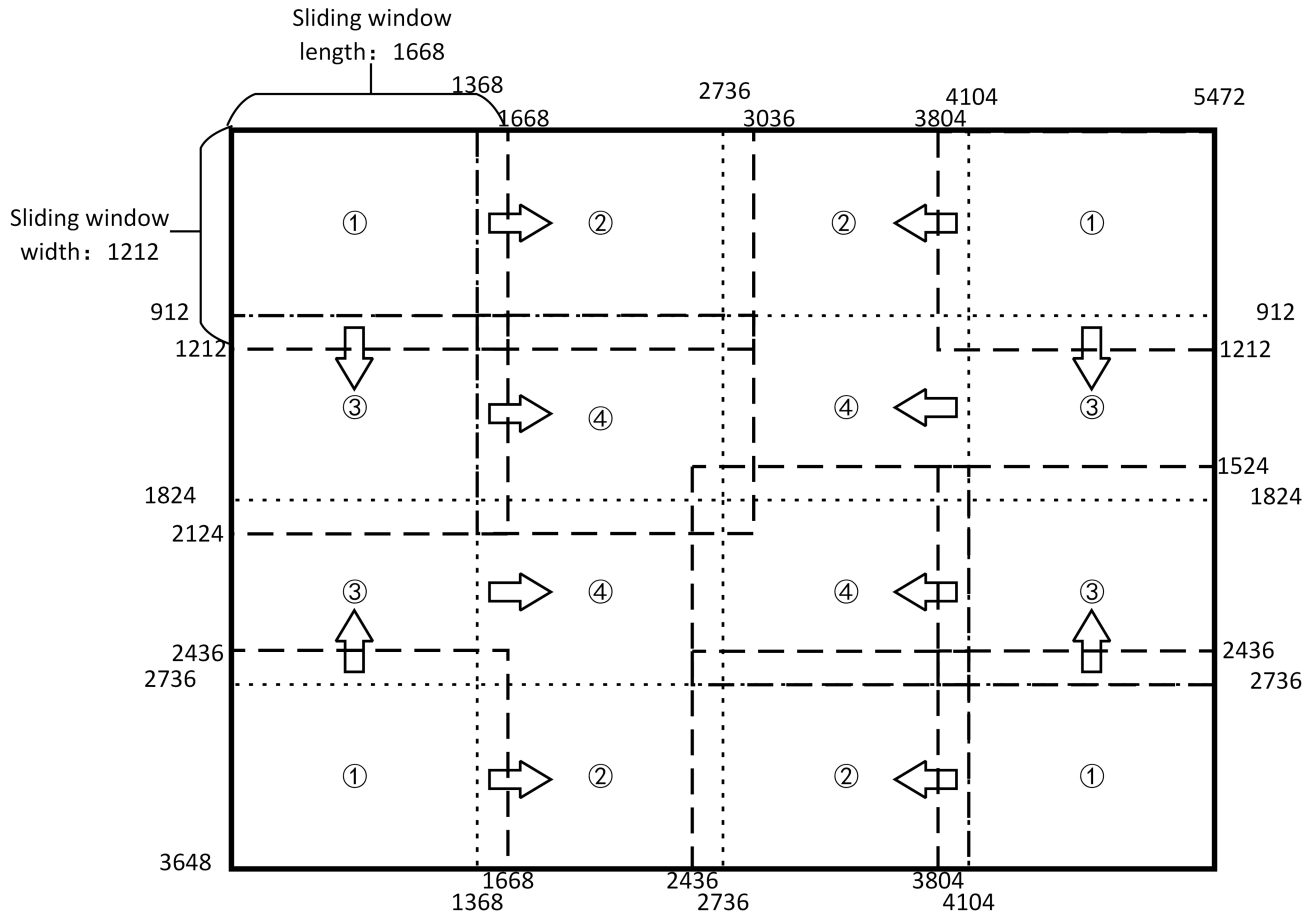


FIGURE 3 | The diagram of overlapping sliding window method.

In our work, the YOLOv4 is used to detect three rice migratory pests from our intelligent searchlight trap based on machine vision. The YOLOv4 model consists of a feature extraction network CSPDarknet-53 and an up-sampling feature fusion module (Bochkovskiy et al., 2020). The activation function for DarknetConv2D of YOLOv4 is Mish and the convolution block is DarknetConv2D_BN_Mish. This design makes it not completely truncate at negative value, thereby ensuring information flow and avoiding the problem of saturation. YOLOv4 uses the CSPnet structure to enhance learning ability through repeated feature extraction. The SPP structure is added to the feature extraction network of YOLOv4, which can greatly increase the receptive field and isolate the most significant contextual features. These improvements enable YOLOv4 to achieve better detection results while consuming less computational resources. The migratory pest targets in this paper have the characteristics of large insect quantity, many insect species, small targets, and similar insects, which put forward higher requirements for the robustness and computational performance of the model. Accordingly, we chose YOLOv4 as the original detection model.

Due to the complex background of migratory pests caused by lots of non-target insects trapped by searchlight trap, the original YOLOv4 model has two detection problems. One is the false detection of target pests and interference pests. The other is the missing detection of small target pests. Aiming at the two problems, we firstly use the overlapping sliding window method to increase the area proportion of the targets in one image. Secondly, the SENet channel attention mechanism is added to the YOLOv4 model to reduce the false detection of target pests. The improved model is named YOLO-MPNet and its network framework is shown in **Figure 4**.

By adding the SENet channel attention mechanism, we design one dependency model of each channel. This model improves the expression ability of the network and makes the network selectively learn some features. Besides, this new network structure can adaptively detect targets by slightly increasing model complexity and a small amount of computation. The specific steps are as follows: (1) to perform global average pooling on the feature layer of the input module, (2) to add two fully connected neural networks and conduct normalization.

Feature Extraction Network

Feature extraction is an important part of target detection. The number of target detection frames, classification accuracy, and the detection efficiency is directly affected by the feature extraction network. The backbone network of the YOLOv4 is CSPDarkNet53,

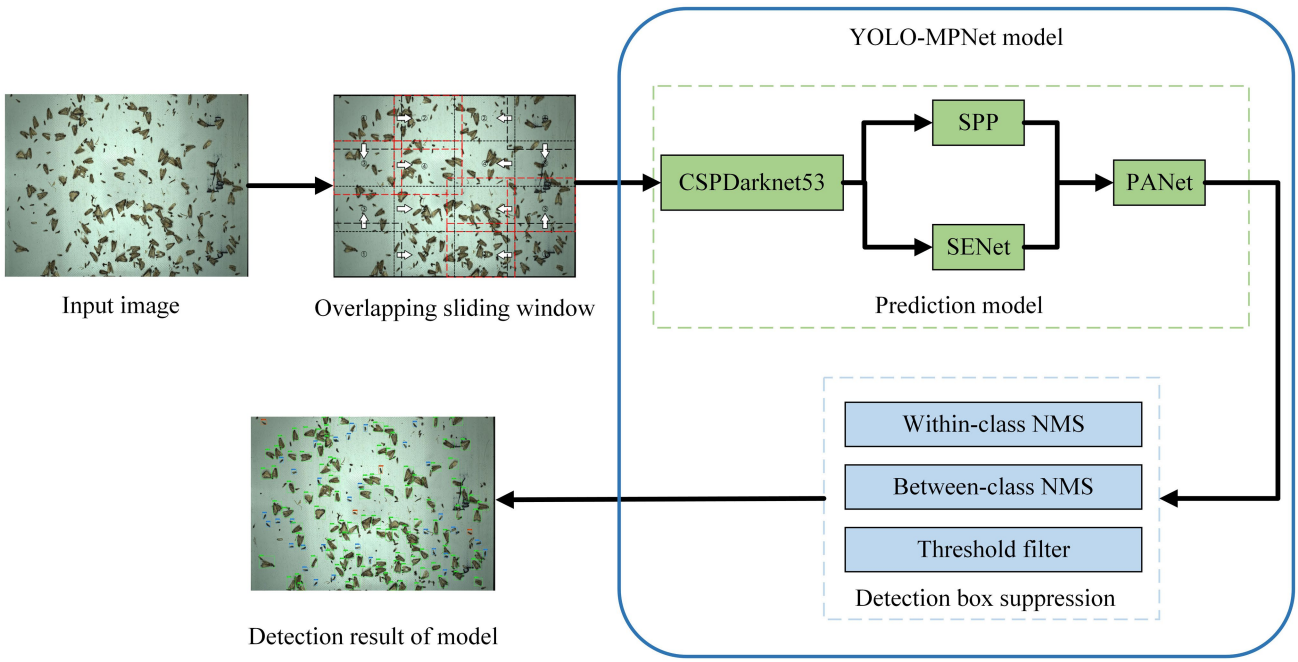


FIGURE 4 | The network framework of YOLO-MPNet model.

which is composed of the resblock body module, one-time down-sampling, and multiple stacking of residual structures.

Initially, input images enter ResBlock by a 3×3 convolution channel. Then, the feature map undergoes multiple down-sampling, which is divided into two 1×1 convolution layers with stride 1 and enters the partial transition and residual block, respectively. After splicing, the feature map finally passed through convolution to reduce the complexity of calculation and improve the calculation speed.

Other Compared Models

To compare the detection performance of different CNN models, we trained another three state-of-the-art target detection models YOLOv3, YOLOv5, and Faster R-CNN.

YOLO model was first proposed by Redmon in 2016. YOLOv3 is the third iterating version and has a great improvement on the detection accuracy and speed (Redmon and Farhadi, 2018). YOLOv5 introduced multi-scale network detection to furtherly enhance the model flexibility (Glenn, 2020). Faster R-CNN is a two-stage target detection model, which combines the candidate region generation stage with the classification stage, and can achieve a high detection accuracy (Girshick, 2015).

Model Training

All models run on a PC with an Intel Core i7-9800xCPU @ 3.8 GHz and 3 GeForce GTX 1080Ti. The operating system is Linux16.04. YOLO series model and Faster R-CNN model run on tensorflow framework.

Evaluation Metrics

To evaluate the detection effect of the YOLO-MPNet model, we use precision (P), recall (R), and F_1 as evaluation indicators. Precision indicates the proportion of the target pests that are correctly detected among all detected targets. Recall indicates the proportion of correctly detected pests among the target pests. F_1 is a comprehensive evaluation index of precision and recall, which is used to evaluate model performance when precision and recall are in conflict. The higher the F_1 value, the better the balance of precision and recall. The formulas are as follows.

$$P = \frac{\text{number of correctly detected pests}}{\text{total number of detected targets}} \tag{1}$$

$$R = \frac{\text{number of correctly detected pests}}{\text{total number of target pests}} \tag{2}$$

$$F_1 = 2\times\frac{P\times R}{P+R} \tag{3}$$

In our work, the pest detection speed is very important in pest occurring peaks. To evaluate the detection speed of different models, frames per second (FPS) is calculated.

RESULTS

Detection Results of Different Models

Table 2 presents the detection results of three migratory pests on the same test set using YOlOv3, YOLOv4, Yolov5, Faster R-CNN, YOLOv4 with OSW and YOLO-MPNet with OSW.

YOLOv4 achieves the higher precision rate, recall rate, and F_1 of three pests than YOLOv3, YOLOv5 and Faster R-CNN. The precision rate of *C. medinalisare*, *S. furcifera*, and *N. lugens*. are 73.66, 66.72, and 71.27%, respectively, and their recall rate are 60.13, 55.24, and 59.24%, respectively. In our mind, the two-stage model Faster R-CNN should have higher precision than one-stage model YOLOv4. As it can be seen, Faster R-CNN seems to be an unsatisfied approach in our pest detection task. Although the FPS of YOLOv5 is higher than YOLOv4, we consider both the precision rate and FPS. So we select the YOLOv4 as an original model which is improved.

The YOLOv4 with overlapping sliding window method effectively improves the precision and recall rates of three pests. The precision rates of *C. medinalisare*, *S. furcifera*, and *N. lugens* are increased by 7.92, 9.7, and 8.19% respectively, their recall rate are increased by 22.26, 19.39, and 18.42%, respectively. Because the sliding window processing method during image preprocessing increases the area ratio of each target pest in the subimages, which helps to extract more abundant features of small target pests and reduce the missing detection.

The searchlight trap attracts a large number of non-target insects. Some insects are similar to target pests, which results in false detection. The improved model YOLO-MPNet with a SENet attention mechanism achieves better detection effects of three migratory pests than YOLOv4 after the same sliding window method is processed on original images. The precision rates of three pests of *C. medinalisare*, *S. furcifera*, and *N. lugens* are increased by 12.56, 9.4, and 9.33% respectively, their recall rates are increased by 9.6, 7.84, and 7.34%, respectively. It proves that the SENet channel attention mechanism can effectively decrease false detection between target pests and interference pests.

Precision-Recall Analysis

To investigate the false detections and missing detections, PR curves of YOLOv4, YOLOv4 with overlapping sliding window and our improved model YOLO-MPNet with overlapping sliding window are shown in **Figure 5**. When pest images are processed with overlapping sliding window, the precisions of YOLOv4 and YOLO-MPNet can keep a high value in a big range of recall. So the overlapping sliding window method can effectively

TABLE 2 | The detection results of different models for migratory pests.

Detection models	Precision (%)			Recall (%)			F_1 (%)			FPS
	C. medinalis	*S. furcifera*	*N. lugens*	*C. medinalis*	*S. furcifera*	*N. lugens*	*C. medinalis*	*S. furcifera*	*N. lugens*	
YOLOv3	70.13	60.69	64.51	57.36	45.13	48.65	63.10	51.76	55.47	0.89
YOLOv4	73.66	66.72	71.27	60.13	55.24	59.24	66.21	60.44	64.70	0.95
YOLOv5	71.26	63.54	69.22	59.21	52.51	56.32	64.68	57.50	62.11	1.02
Faster R-CNN	72.21	54.61	55.76	59.68	47.23	48.11	65.35	50.65	51.65	0.32
YOLOv4 with OSW	81.58	76.42	79.46	82.39	74.63	77.66	81.98	75.51	78.55	0.68
YOLO-MPNet with OSW	94.14	85.82	88.79	91.99	82.47	85.00	93.05	84.11	86.85	0.66

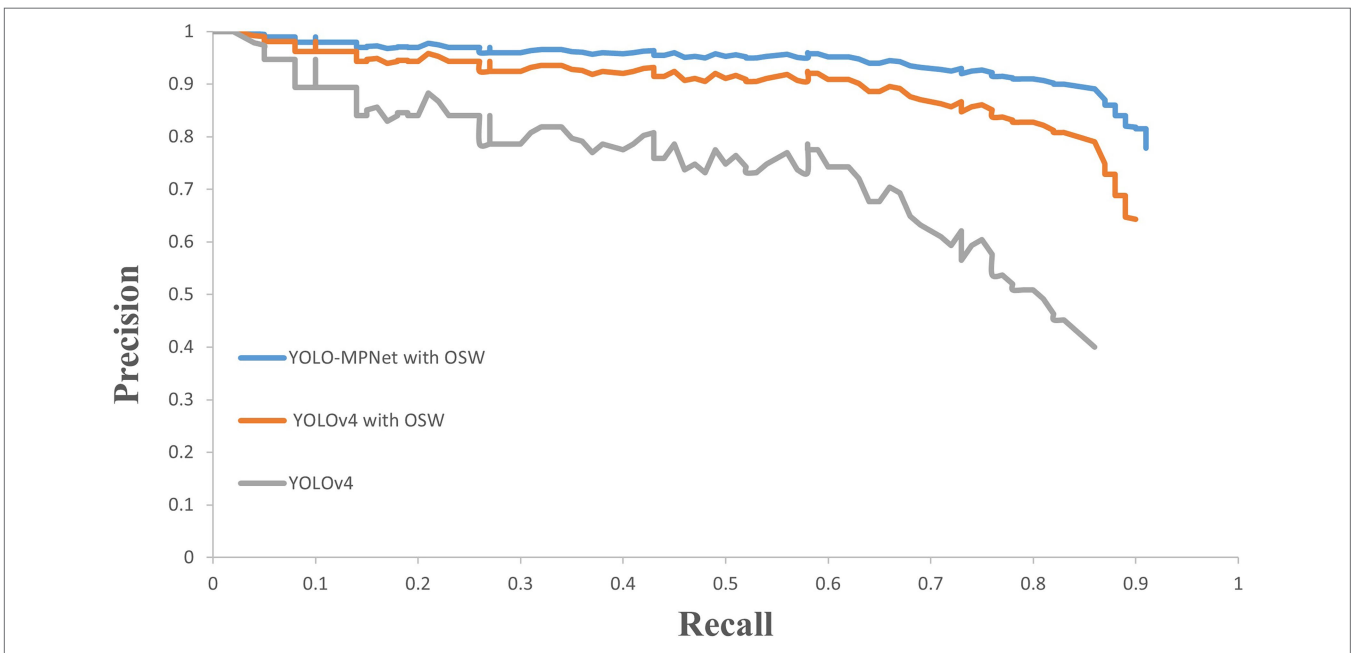


FIGURE 5 | PR curves for different models.

reduce pest false detections and missing detections. In general, YOLO-MPNet performs the best on three migratory pest detection with a high precision and recall at same time.

Visualization of Detection Results

The detection results of migratory pests are visualized in **Figure 6**. YOLO-MPNet could detect the three migratory pests well under different insect densities. As it can be seen, the trapped pests could effectively be dispersed by our multilayer insect conveyor belts. Some of occluded pests could be correctly detected by our model.

Web Client Interface of System

The web client interface of the intelligent monitoring system of migratory pests mainly includes user login, automatic identification of migratory pests, equipment management, user

FIGURE 6 | Examples of detected pests. The green, orange, and blue boxes contain *C. medinalis*, *S. furcifera*, and *N. lugens*, respectively.

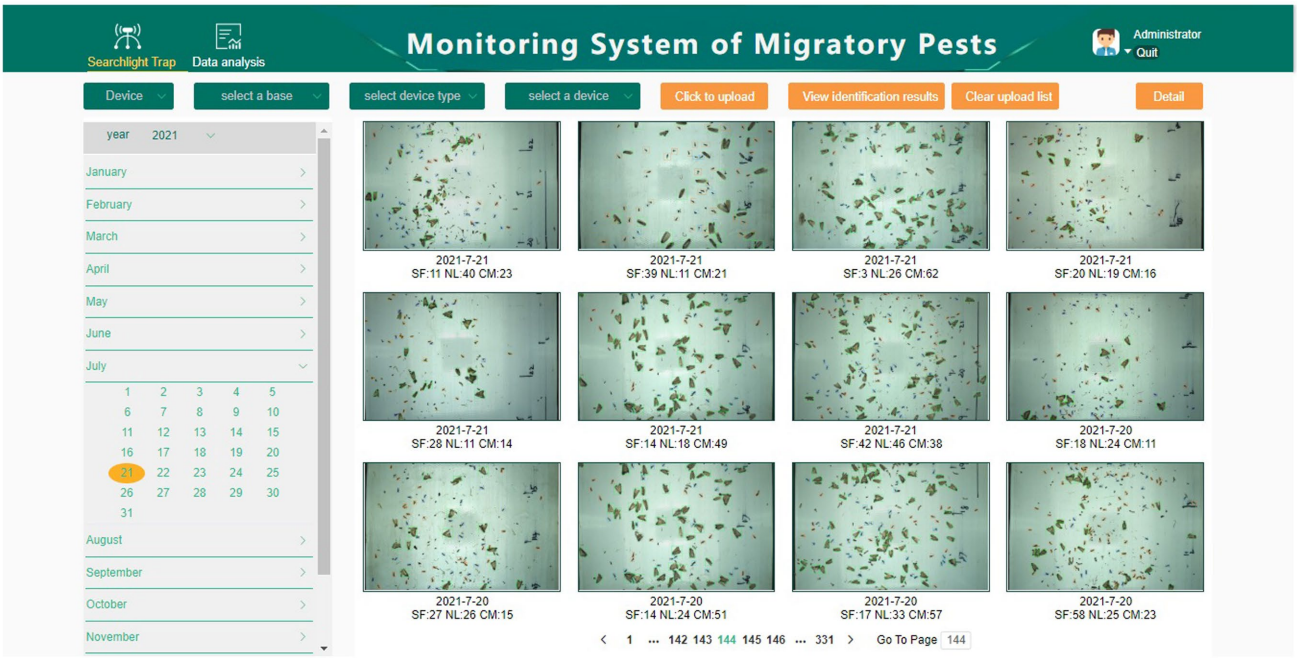


FIGURE 7 | Web client interface of intelligent monitoring system of migratory pests.

management, and data curves. Users can view the detection result images through the web interface and historical monitoring data. **Figure 7** shows the web interface of the system and the detected images.

CONCLUSION

To realize automatic and accurate identification of rice migratory pests from searchlight traps, we develop an intelligent monitoring system of migratory pests, which is composed of a searchlight trap based on machine vision, an automatic identification model of migratory pests, a Web client, and a cloud server. To identify and count three rice migratory pests (*C. medinalis, S. furcifera*, and *N. lugens*) from a large number of non-target insects trapped by searchlight traps, we propose an improved model, YOLO-MPNet. To solve the problem that the backbone network cannot effectively extract features of small target pests, this paper introduces the overlapping sliding window processing method, which can improve the area proportion of small targets in images and optimize the identification effect of small target pests. At the same time, the feature extraction network is improved by adding the SENet channel attention mechanism. The model's adaptability to complex backgrounds is strengthened. YOLO-MPNet has achieved higher precision, recall and F_1 values among three species of migratory pests (*C. medinalis*, *S. furcifera*, and *N. lugens*) than the YOLOv3, YOLOv4, YOLOv5, and Faster R-CNN models.

In this paper, only three species of rice migratory pests are identified by our model. In fact, some non-migratory pests are trapped by the searchlight traps. In future work, more species of pests from searchlight traps will be studied.

DATA AVAILABILITY STATEMENT

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

AUTHOR CONTRIBUTIONS

GS, HL, and QY proposed the detection model. GS, HL, SL, QY, and JX wrote and revised the manuscript. SL, BY, JL, and JT contributed to paddy fields and manual data annotation. GS, HL, and ZF developed the system software. All authors contributed to the article and approved the submitted version.

FUNDING

This study is supported by the National Key Research Program of China during the 14th Five-Year Plan Period (no. 2021YFD1401100) and the Natural Science Foundation of Zhejiang, China (no. LY20C140008).

ACKNOWLEDGMENTS

The authors would like to thank China National Rice Research Institute for providing the experiment paddy fields.

REFERENCES

Bochkovskiy, A., Wang, C.-Y., and Liao, H.-Y. M. (2020). Yolov4: Optimal speed and accuracy of object detection. arXiv [Preprint]. doi: 10.48550/arXiv.2004.10934

Feng, H. (2003). Community Aloft and Radar Observations of Seasonal Migration of Insects in Northern China. Doctor's Thesis. Beijing: Chinese Academy of Agricultural Sciences.

Feng, H. (2011). Application of radar in entomological research. *Plant Prot.* 37, 1–13. doi: 10.3969/j.issn.0529-1542.2011.05.001

Feng, J. (2020). Research and Optimization on Rice Light-Trap Pest Detection Method Based on Deep Learning. Master's Thesis. Hangzhou: Zhejiang Sci-Tech University.

Girshick, R. (2015). "Fast r-cnn." in *Proceedings of the IEEE International Conference on Computer Vision*. December 7–13, 2015; 1440–1448.

Glenn, J. (2020). YOLOv5 [Online]. Available at: https://github.com/ultralytics/yolov5 (Accessed February 20, 2022).

Han, R., and He, Y. (2013). Remote automatic identification system of field pests based on computer vision. *Trans. Chin. Soc. Agric. Eng.* 29, 156–162. doi: 10.3969/j.issn.1002-6819.2013.03.021

Hu, G., Gao, B., Feng, H., Jiang, X., Zhai, B., and Wu, K. (2020). Insect migration: individual behaviour, population dynamics and ecological consequences. *Sci. Found. Chin.* 34, 456–463. doi: 10.16262/j.cnki.1000-8217.2020.04.011

Jiang, Y., Liu, J., and Ceng, J. (2016). Using vertical-pointing searchlight-traps to monitor population dynamics of the armyworm Mythimna separate (Walker) in China. *Chin. J. Appl. Entomol.* 53, 191–199. doi: 10.7679/j.issn.2095-1353.2016.024

Jiang, Y., Liu, J., Ceng, J., Huang, C., and Zhang, T. (2021). Occurrence of, and damage caused by, major migratory pests and techniques for monitoring and forecasting these in China. *Chin. J. Appl. Entomol.* 58, 542–551. doi: 10.7679/j.issn.2095-1353.2021.056

Jiao, G., Ceng, J., Shang, X., and Liu, J. (2017). High altitude monitoring lamp and automatic pest monitoring lamp comparative experiment of monitoring slime worms. *Chin. Plant Prot.* 37, 41–44.

Lee, J.-S. (1980). Digital image enhancement and noise filtering by use of local statistics. *IEEE Trans. Pattern Anal. Mach. Intell.* 2, 165–168. doi: 10.1109/TPAMI.1980.4766994

Liu, Z., Gao, J., Yang, G., Zhang, H., and He, Y. (2016). Localization and classification of paddy field pests using a saliency map and deep convolutional neural network. *Sci. Rep.* 6:20410. doi: 10.1038/srep20410

Qin, B. (2019). Comparative analysis of the effect of two indicators in green prevention and control. *South Chin. Agric.* 13, 136–138. doi: 10.19415/j.cnki.1673-890x.2019.14.071

Qiu, D., Zhang, H., Liu, X., and Liu, Y. (2007). Design of detection system for agriculture field pests based on machine vision. *Trans. Chin. Soc. Agric. Mach.* 38, 120–122.

Redmon, J., Divvala, S., Girshick, R., and Farhadi, A. (2016). "You only look once: Unified, real-time object detection." in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*. June 27–30, 2016; 779–788.

Redmon, J., and Farhadi, A. (2017). "YOLO9000: better, faster, stronger." in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*. July 21–26, 2017; 7263–7271.

Redmon, J., and Farhadi, A. (2018). Yolov3: an incremental improvement. arXiv [Preprint]. doi: 10.48550/arXiv.1804.02767

Shang, X. (2017). Comparative analysis of high altitude lamp and automatic pest monitoring lamp in 2015. *Primary Agric. Technol. Exten.* 5, 58–60.

Song, H., Li, L., Zhang, Q., Sun, C., Li, C., Lu, Z., et al. (2021). The species and population dynamics of insects attracted by searchlight traps in Jinan Shandong. *J. Plant Prot.* 48, 927–928. doi: 10.13802/j.cnki.zwbhxb.2021.2020164

Yan, C., Ruan, X., Wang, W., Hu, L., Li, X., Wang, Y., et al. (2021). Analysis of insect population community structure under Xinyang's serachlight. *J. Xinyang Agric. For. Univ.* 3, 107–110. doi: 10.16593/j.cnki.41-1433/s.2021.03.023

Yang, R., Zhu, X., and Zhu, F. (2017). Research and development application process and development of crop disease and pest monitoring and investigation tools in China. *Chin. Plant Prot.* 37, 51–55. doi: 10.3969/j.issn.1672-6820.2017.01.012

Yao, Q., Yao, B., Lu, J., Tang, J., Feng, J., and Zhu, X. (2021b). Research on fine-grained image recognition of agricultural light-trap pests based on bilinear attention network. *Sci. Agric. Sin.* 54, 4562–4572. doi: 10.3864/j.issn.0578-1752.2021.21.007

Yao, Q., Wu, S., Kuai, N., Yang, B., Tang, J., Feng, J., et al. (2021a). Automatic detection of rice planthoppers through light-trap insect images using improved CornerNet. *Trans. Chin. Soc. Agric. Eng.* 37, 183–189. doi: 10.11975/j.issn.1002-6819.2021.07.022

Zhang, Z., Zhang, Y., Jiang, Y., Zhang, L., and Cheng, D. (2017). Development of radar entomology and related prospects for future application. *Plant Prot.* 43, 18–26. doi: 10.3969/j.issn.0529-1542.2017.05.003

Zou, X. (2013). Research on field identification technology of rice planthopper based on machine vison. Doctor's Thesis. Nanjing: Nanjing Agricultural University.

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

frontiers | Frontiers in Plant Science

ORIGINAL RESEARCH
published: 07 July 2022
doi: 10.3389/fpls.2022.908114

Promoting the Development of *Astragalus mongholicus* Bunge Industry in Guyang County (China) Based on MaxEnt and Remote Sensing

Ru Zhang[1,2], Mingxu Zhang[1], Yumei Yan[1], Yuan Chen[3], Linlin Jiang[3], Xinxin Wei[4], Xiaobo Zhang[5], Huanting Li[1] and Minhui Li[1,2,4,5,6]**

[1]Baotou Medical College, Baotou, China, [2]Inner Mongolia Hospital of Traditional Chinese Medicine, Hohhot, China, [3]Inner Mongolia Key Laboratory of Characteristic Geoherbs Resources Protection and Utilization, Baotou, China, [4]Department of Pharmacy, Inner Mongolia Medical University, Hohhot, China, [5]School of Life Sciences, Inner Mongolia University, Hohhot, China, [6]State Key Laboratory Breeding Base of Dao-di Herbs, National Resource Center for Chinese Materia Medica, China Academy of Chinese Medical Sciences, Beijing, China

To provide high-quality *Astragalus mongholicus* Bunge to domestic and foreign markets and maintain sustainable development of the *A. mongholicus* industry, Firstly, we evaluated the impact of environmental factors and planting areas on the *A. mongholicus* industry. The maximum entropy method (MaxEnt) was utilized to simulate the suitability distribution of *A. mongholicus* and establish the relationship between the active component contents of *A. mongholicus* and ecological factors through linear regression analysis. The random forest algorithm was subsequently used to perform feature selection and classification extraction on Sentinel-2 imagery covering the study area. Furthermore, the planting, processing, and sales of *A. mongholicus* in Guyang County were investigated, and the roles of stakeholders in the value chains were analyzed. The results demonstrated that precipitation of the warmest quarter, minimum temperature of the coldest month, standard deviation of seasonal temperature changes, range of mean annual temperature, and mean diurnal range [mean of monthly (max temp - min temp)] were the five environmental variables that contributed the most to the growth of *A. mongholicus*. The most influential factor on the distribution of high-quality *A. mongholicus* was the mean temperature of the coldest quarter. The classification results of image features showed that the planting areas of *A. mongholicus* was consistent with the suitable planting areas predicted by MaxEnt, which can provide data support to the relevant departments for the macro development of the *A. mongholicus* industry. In the production of *A. mongholicus*, 10 value chains were constructed, and the study demonstrated that the behavior of stakeholders, target markets, and the selected planting area had a significant impact on the quality of *A. mongholicus*.

Keywords: *Astragalus mongholicus* Bunge, value chains, species distribution models, maximum entropy, remote sensing, random forest

OPEN ACCESS

Edited by:
Lei Shu,
Nanjing Agricultural University, China

Reviewed by:
Le Li,
Cornell University, United States
Xingfeng Chen,
Aerospace Information Research Institute (CAS), China
Chunnian He,
Chinese Academy of Medical Sciences and Peking Union Medical College, China
Jihua Meng,
Aerospace Information Research Institute (CAS), China

***Correspondence:**
Huanting Li
lihuanting2021@163.com
Minhui Li
prof_liminhui@yeah.net

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

Received: *30 March 2022*
Accepted: *15 June 2022*
Published: *07 July 2022*

Citation:
Zhang R, Zhang M, Yan Y, Chen Y, Jiang L, Wei X, Zhang X, Li H and Li M (2022) Promoting the Development of Astragalus mongholicus Bunge Industry in Guyang County (China) Based on MaxEnt and Remote Sensing. Front. Plant Sci. 13:908114. doi: 10.3389/fpls.2022.908114

INTRODUCTION

The continuous improvement of people's health awareness has resulted in medicinal plants receiving increased attention. *A. mongholicus* is an important raw material for use in functional health care products and is widely sought after. In China, the root of *A. mongholicus* is a traditional Chinese medicine (**Figure 1**). Moreover, *A. mongholicus* plants are rich in glycosides, polysaccharides, and flavonoids that exert various beneficial health effects, such as immunoregulation, protection against cardiovascular and cerebrovascular diseases, delayed aging, and anti-rheumatism effects (Durazzo et al., 2021).

In recent years, the growth of international and domestic market demand for *A. mongholicus* has prompted the expansion of its plantation areas, in an effort to alleviate the supply and demand pressures. Guyang County is renowned as one of the main producing areas of high-quality *A. mongholicus*. Accordingly, the *A. mongholicus* produced in Guyang County was awarded the "Geographical Indication of Agricultural Products" status by the Ministry of Agriculture and Rural Affairs in 2017.[1] Furthermore, subsequent to its review as a high-quality agricultural product by the working institutions of the prefecture- and provincial-level agricultural and rural departments in September 2020, *A. mongholicus* from Guyang County was included in the second batch of national famous, special, and excellent new agricultural products.[2] Therefore, Guyang county is considered an ideal location for the production of high-quality *A. mongholicus* (Dao-di Herbs).

Due to the rotation cultivation mode of *A. mongholicus*, a large number of suitable land is needed to provide options for planting areas in the coming years. However, to obtain the planting situation of *A. mongholicus*, it also needs sufficient time to consult the data and conduct field investigation. This survey method is expensive, inefficient, and prone to subjective deviation, resulting in errors in planting conditions. Species distribution models (SDMs) use specific algorithms to correlate species distribution with environmental variables, thereby predicting the potential distribution of a species across different geographical spaces and time (Zhu et al., 2013). This probability reaction reflects habitat suitability and allows for further research on the regional distribution of the estimated model of the target species. Among the SDMs, the maximum entropy model (MaxEnt) is most commonly used (Xu et al., 2015) which produces good results. For example, Yang et al. (2020b) successfully used the MaxEnt model to predict the habitat suitability of *Salweenia bouffordiana* by analyzing the main environmental factors affecting its habitat (Yang et al., 2020b). Considering this, using the MaxEnt model to understand the relationship between different environmental factors and species distribution is an effective way to develop management and protection strategies for medicinal plant species.

Remote sensing can play significant roles in determining the distribution, growth area and status, and occurrence of diseases and insect pests among medicinal plants. Since 2000, remote sensing technology has gradually advanced the theory and methods of medicinal plant resource investigations and has been effectively applied in medicinal plant production (Lan et al., 2021). Previous studies have successfully utilized Landsat-8TM, Gaofen-1 (GF-1), Resources satellite three (ZY-3), and other remote sensing satellite data for the detection of medicinal plants (Na et al., 2013; Shi et al., 2017). Sentinel-2 is mainly used for land environmental monitoring and provides information on land cover, including soil condition and vegetation patterns. This information is crucial for improving agricultural and forestry planting structures, estimating agricultural areas, and predicting crop yield. On this basis, using Sentinel-2 imagery to comprehensively monitor the planting distribution of *A. mongholicus* in Guyang County may be of great significance to the development of the *A. mongholicus* industry.

With the exception of planting area, planting personnel, processing, sales and other factors also affect the development of *A. mongholicus* industry. Currently, *A. mongholicus* is mainly cultivated by individual farmers, and only a few large-scale cultivation enterprises exist. These individual farmers are restricted by limited funds, small-scale planting, and lack of cultivation knowledge, and the consequential blind planting results in failure to meet standardized planting requirements. In addition, variety mixing and species degradation in the seed supply base (Qi, 2020), the use of inappropriate planting habitats, and dense application of herbicides, fertilizers, and insecticides may reduce crop quality and be detrimental to

[1]http://www.moa.gov.cn/nybgb/2017/dsq/201802/t20180201_6136210.htm
[2]http://www.aqsc.agri.cn/tzgg/202009/t20200918_361324.htm

FIGURE 1 | The cultivated of *Astragalus mongholicus* **(A)**; Details of *A. mongholicus* **(B)**; dried roots *A. mongholicus* **(C)**.

consumer health (Zhao et al., 2017). In contrast, large-scale planting companies have mainly been able to circumvent these problems by hiring experienced professionals and utilizing specialized mechanical equipment to meet planting standards. Medicinal plant products pass through several stakeholder levels, including processing and distribution, before finally reaching the consumer. The processing of *A. mongholicus* is primarily not very extensive, and is mainly executed in family workshops using relatively outdated technology. Moreover, the foundational processing technologies of some enterprises also requires improvement to avoid possible imbalances in the quality of medicinal materials (Sun and Chen, 2018). The value chain (VC) describes the entire process from the initial planting of raw materials to the processing and sale of final products, all while considering the relationships between different stakeholders in the chain (Booker et al., 2012). In recent years, an increasing number of people are applying the VCs to the Chinese herbal medicine industry. According to the production and circulation mode, the roles of different stakeholders in the VCs are organized to effectively evaluate the quality and economic benefits of Chinese herbal medicine in the different VCs (Booker et al., 2015).

In addition to the utilization of *A. mongholicus* as a traditional Chinese medicine, it has been incorporated into Chinese patent medicine, food therapeutics, and other health care products (Qin et al., 2013), resulting in a gradual expansion of its market demand. The artificial planting of high-quality *A. mongholicus* may be an effective method to bridge the gap between supply and demand. To this end, this study aims to (1) predict the suitable growth area of *A. mongholicus* by utilizing MaxEnt and monitoring ecological factors affecting its production. Linear regression analysis was used to evaluate the relationship between the active component content of *A. mongholicus* and ecological factors, and to identify those ecological factors that contributed most to the growth of *A. mongholicus*; (2) monitor planting distribution areas by combining Sentinel-2 imagery; and (3) visit and inspect the planting, processing, and sales links of *A. mongholicus*, build VCs, and analyze the impact of stakeholder behavior on the development of the *A. mongholicus* industry.

MATERIALS AND METHODS

Study Area Description

Guyang County is located in central Inner Mongolia (**Figure 2**), at latitude 40°42′–41°08′ N and longitude 109°40′–110°41′ E. The county stretches across approximately 80 km (east to west) and is roughly 66 km wide (north to south), covering a total area of 5,025 km^2. Guyang county has a mid-temperate, continental, arid, and semi-arid monsoon climate, with low average temperatures, little precipitation, and sufficient sunlight. Furthermore, major temperature variations are characteristic of Guyang county, and its mountains and hills make up approximately 90% of the total area (Rural Social and Economic Investigation Department of National Bureau of Statistics, 2021). Chunkun Mountain, in the east of the territory, is 2,321 m above sea level, dividing the county into the southern mountainous and northern hilly areas. The area in between constitutes the Guyang and Bailingnuo basins, of which the lowest point is 1,240 m above sea level. Guyang presents with a typical "plateau basin" terrain that performs an adequate heat collection function, which is beneficial to the growth of *A. mongholicus*.

Study Species and Data Collection

In this study, the sample data were derived from a field survey of Guyang county from July 8 to July 12, 2021 using the Global Positioning System (GPS), from which the geographic distribution data of *A. mongholicus* were obtained. During the positioning process, sample points were randomly selected from each plot subject to the condition that the minimum distance between any two points would be at least 20 m. Subsequently, from 8 October to 12 October 2021, samples of mature *A. mongholicus* from Guyang County were collected for active ingredient content determination. Different stakeholders of the *A. mongholicus* industry in Guyang County were interviewed, and relevant information on the cultivation, processing, and sales stages of *A. mongholicus* was obtained using field investigations.

MaxEnt Predicted *Astragalus mongholicus* Distribution

Climate is the main environmental factor determining species distribution. The MaxEnt model was used to explore the environmental niche and potential distribution of *A. mongholicus* according to its spatial location and environmental variables, thereby allowing for the selection of a suitable planting area for *A. mongholicus*.

Ecological Factor Selection

A total of 19 climate (with a resolution of 30 s) and altitude datasets were obtained from the global climate data website.[3] We selected 20 ecological factors, including temperature, precipitation, and altitude, to determine which ecological factors have the greatest impact on the development of the planting industry (Ngarega et al., 2021; Zhang et al., 2021).

MaxEnt Model Prediction

Jaynes et al. proposed the MaxEnt principle in 1957 (Jaynes, 1957), and Phillips et al. developed MaxEnt software—that is easy to operate and does calculations swiftly—using the Java™ programming language (Phillips et al., 2006), based on the original model. Herein, MaxEnt software (Version 3.4.1 K) was used to predict the suitable distribution of *A. mongholicus*, and 10-fold cross-validation was conducted to analyze the validity and accuracy of the model. The jackknife method was used to measure the weight of each variable and output it in logistic format. The maximum number of iterations and the convergence domain were set to 10^5 and 0.0005, respectively. The operation was repeated 10 times, and the remaining values were used as default values to extract and analyze habitat

[3]http://www.worldclim.org/

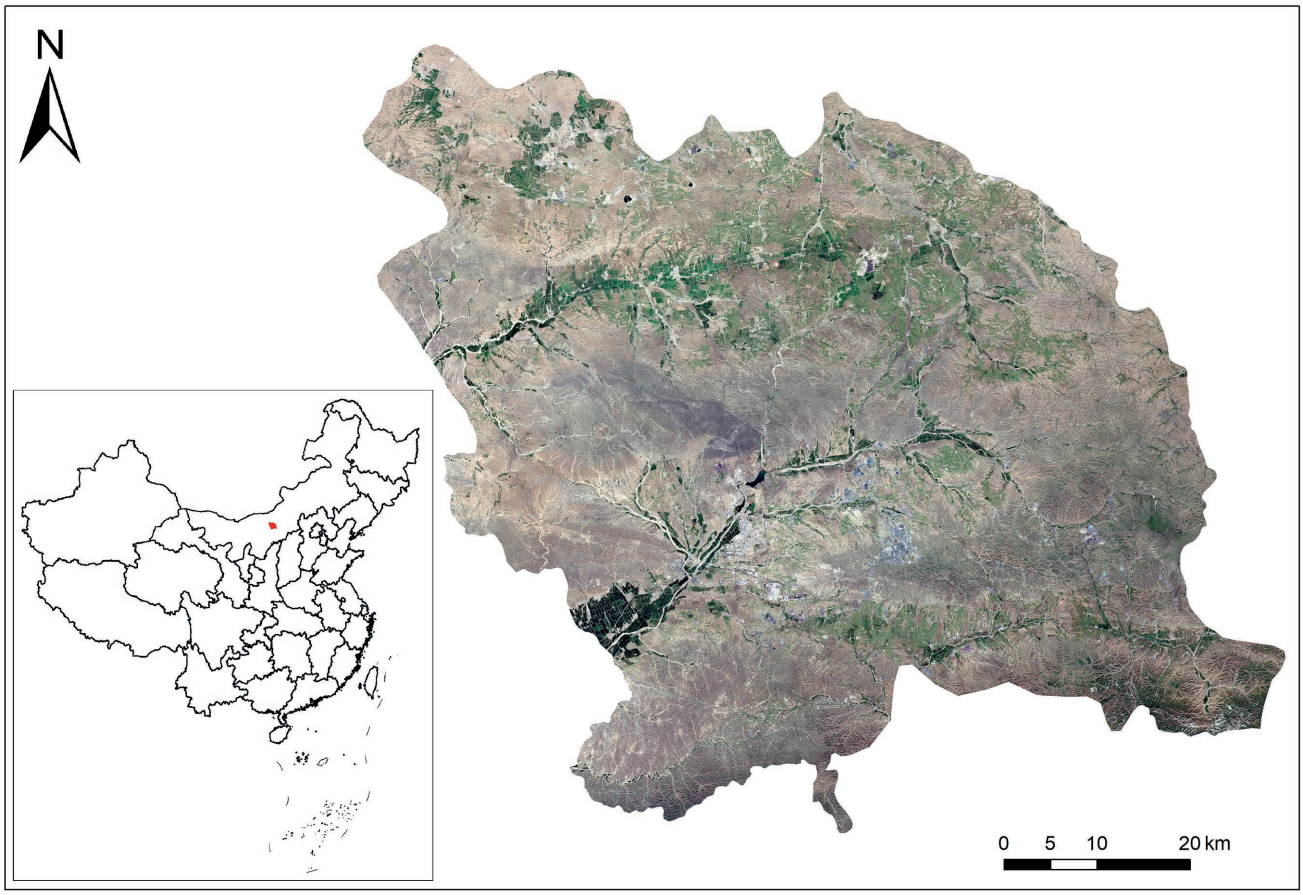


FIGURE 2 | Land cover map of Guyang County.

suitability. Using the latitudinal and longitudinal geographical distribution data of *A. mongholicus* and relevant ecological factor data from the study area, a prediction model was established for the potential geographical distribution of *A. mongholicus* (Zhu et al., 2017). The model results were evaluated using the area under the receiver operating characteristic (ROC) curve (area under curve, AUC). The AUC-based evaluation standard indicated whether the prediction result was less accurate (<0.5), acceptable (0.5–0.8), or ideal (0.8–0.9), the latter of which would demonstrate high modeling accuracy.

Construction of Active Component Content and Its Relationship With Main Ecological Factors

Saponins and flavonoids are the main active components in *A. mongholicus*, which serve as valuable indices for evaluating the quality of *A. mongholicus* in Chinese, British, and European pharmacopoeia. According to the "Chinese Pharmacopoeia" (2020 edition), the astragaloside IV and calycosin-7-glucoside content in 37 *A. mongholicus* samples had been determined using high-performance liquid chromatography (HPLC; European Pharmacopoeia Commission, 2017; British Pharmacopoeia Commission, 2019; National Pharmacopoeia Committee, 2020). In the present study, astragaloside IV was determined using the Thermo Fisher Ultimate 3,000 HPLC system, Agilent C18 column, with a flow rate of 1.0 ml/min. The mobile phase consisted of acetonitrile and deionized water (34%:66%), and the injection volume was 10 μl. Furthermore, calycosin-7-glucoside was separated on a Waters C18 column at a flow rate of 1 ml/min. The mobile phase consisted of acetonitrile (solvent A) and water containing 0.2% methanoic acid (solvent B). Gradient elution was applied as follows: 0–20 min, 80–60% B; 20–30 min, 60% B. The injection volume was 10 μl and the temperature of the column was maintained at 30°C. Detection was performed at a wavelength of 260 nm and each sample was assayed in triplicate.

In addition, SPSS statistical analysis software was utilized to analyze the differences in astragaloside IV and calycosin-7-glucoside contents in *A. mongholicus* from different township areas in Guyang County. A correlation matrix was used to determine the relationship between astragaloside IV, calycosin-7-glucoside, and the major ecological factors affecting *A. mongholicus* growth. By conducting stepwise linear regression analysis, the relationship equations between each index component and the main ecological factors were obtained.

The relationship equations were input into the grid calculator of ArcGIS to obtain the quantitative distribution layers of

TABLE 1 | Spectral characteristic index set.

Spectral region	Vegetation index	Formula
Normalized vegetation index	NDVI	(B8-B4)/(B8 + B4)
Enhanced vegetation index	EVI	2.5 × (B8-B4)/ (B2 + 6 × B1 + 7.5 + B3 + 1)
Ratio vegetation index	RVI	B8/B4
Red edge vegetation index 1	NDVIre1	(B8 – B5)/(B8 + B5)
Normalized difference vegetation index red-edge 1narrow	NDVIre 1n	(B8A – B5)/(B8A + B5)
Normalized difference vegetation index red-edge 2	NDVIre 2	(B8 – B6)/(B8 + B6)
Normalized difference vegetation index red-edge 2narrow	NDVIre 2n	(B8A – B6)/(B8A + B6)
Normalized difference vegetation index red-edge 3	NDVIre 3	(B8 – B7)/(B8 + B7)
Normalized difference vegetation index red-edge 3narrow	NDVIre 3n	(B8A – B7)/(B8A + B7)

astragaloside IV and calycoside-7-glucoside in *A. mongholicus*. Using the spatial calculation function of ArcGIS, these two layers were superimposed on the ecological suitability distribution layer of *A. mongholicus*, from which the spatial suitability distribution region of astragaloside IV and calycoside-7-glucoside in Guyang County was obtained.

Acquisition and Processing of Remote Sensing Data

Image Data Acquisition

In recent years, the continuous development of remote sensing technology and its applications has greatly reduced the human error of traditional manual field surveillance, improved the objectivity, scientificity, and accuracy of survey data, and has been widely used to survey medicinal plant resources (Chen, 2021). We aimed to use remote sensing technology to monitor the planting area of *A. mongholicus*, obtain data on its planting practices and area, provide references for relevant departments to formulate related policies and plans, ensure product supply and stable market prices, and promote the steady development of the *A. mongholicus* industry.

The Sentinel-2 satellite was launched by the European Space Agency (ESA) for the EU Copernicus Programme to support global land services, including the monitoring of vegetation, soil and water coverage, inland waterways, and coastal areas. Sentinel-2 imagery cover 13 spectral bands of visible light. Moreover, considering the good spatial resolution, global coverage, and relatively good time resolution of Sentinel-2 data, it is widely used in many fields.

Images covering the entire Guyang County (February–November 2021) were downloaded from the Copernicus Open Access Center.[4] Although some associated data have been orthogonal and geometric corrected, they cannot be directly used. As such, preprocessing, such as atmospheric correction and resampling, was required and the resolution of all bands was resampled to 10 meters. The pre-processed image was synthesized by the layer stacking function of ENVI. Additionally, bands 2, 3, 4, 5, 6, 7, 8, 8A, 11, and 12 were mainly included, whereafter regional cropping was performed using the subset data of the ROIs function to obtain a complete image of Guyang County.

Time Window Selection for Extraction of the Vegetation Index

According to phenological knowledge, the appropriate time window is selected to extract the characteristics of crop planting structure during the process of crop development. To ensure economic benefits, crop phenological indicators were recorded through field observations. Most *A. mongholicus* in Guyang County were selected and sown in April–May, matured in July–August, and harvested in October–November. Various crops, in addition to *A. mongholicus*, are planted in Guyang County. To allow for better classifications, we extracted the normalized difference vegetation index (NDVI) of Guyang County from February to October 2021 using ENVI software and combined the training datasets of *A. mongholicus* to develop the temporal NDVI profile, thereby obtaining the spectral changes and time windows over different periods (Inoue and Olioso, 2006). The NDVI is used to detect the growth state of vegetation and vegetation coverage. Moreover, it can reflect the background effects of the vegetation canopy, such as soil, wetland, snow, dead leaves, and roughness, etc. In addition, we also extracted the ratio vegetation index (RVI), enhanced vegetation index (EVI), and normalized red edge vegetation index (NDVIre; **Table 1**; Tan et al., 2018).

Principal Component Analysis

Principal component analysis (PCA) was performed on the Sentinel-2 imagery within the optimal time window. PCA is a commonly used dimensionality reduction method in image processing. In remote sensing image classification, PCA is often used to eliminate the correlation between bands and perform feature selection. Under the premise of not reducing the "effective" information, the original dataset was converted into "effective" information, and the identification was carried out with fewer dimensions, thereby reducing the number and dimension (Hess and Hess, 2018; Wang et al., 2021a).

Texture Feature Extraction

Texture features were extracted from the bands containing the main features, based on the PCA results. Texture reflects a certain change rule for an object's surface color and gray level. This information effectively distinguishes ground objects with similar spectra and different spatial distribution structure characteristics, and is widely used to extract image information. Texture features are also important in the process of feature extraction, representing the spatial distribution of pixels in remote sensing images. The addition of texture features can aid in the reduction of the salt and pepper noise and improve classification accuracy (Liu et al., 2021).

[4]https://scihub.copernicus.eu/

Classifier Classification

The random forest classifier (Breiman, 2001), which has the important function of analyzing the significance of features and building a classification regression tree, was used. The collected coordinate points were processed using ArcGIS to generate a sample set, and 70% of the sample points were randomly selected and imported into EnMAP and ENVI software together with remote sensing images as training sets.

Accuracy Evaluation

A confusion matrix was used to classify and identify the planting area of *A. mongholicus*. The matrix consisted of the following terms: overall accuracy (OA), user accuracy (UA), and the Kappa coefficient of variation (Kappa; Vasileios et al., 2018). The overall accuracy represents the probability that the classified result is consistent with the test data type for each random sample. Furthermore, Kappa coefficient is an indicator for comprehensively evaluating the classification accuracy and is used to judge the consistency of images.

VCs Analysis

First, we determined the main links in the production process of *A. mongholicus*, whereafter the stakeholders in the process were identified and matched to different chains, linking the main production activities with the stakeholders according to their different interests. A VC analysis diagram of the roles played by the stakeholders in different chains was then created. In addition, the production behavior, quality, and financial performance of *A. mongholicus* in the VCs were analyzed according to the methods described by Yao et al. (2018) Finally, we analyzed the strengths and weaknesses of the VCs in terms of safety, quality, and geographical indications (Yao et al., 2018).

Price Forecasting

Fluctuations in the price of Chinese herbal medicines not only have a major impact on stakeholders in the VCs, but also exert pressure on governmental market regulation, affecting the sustainability and healthy development of the Chinese medicine industry (Cui et al., 2020). Therefore, the price forecast of Chinese herbal medicines could provide a price reference for stakeholders in the VCs and ensure smooth progression of Chinese herbal medicines planting and production processes. The autoregressive integrated moving average model (ARIMA) is a time-series autoregressive technique that calculates future short-term forecasts by analyzing historical data (Alabdulrazzaq et al., 2021). It was created by Box and Jenkins in the 1970s to mathematically describe changes in a time series (Rao et al., 1972). In the current study, price data of *A. mongholicus* from 2017 to 2021 was collected and the future market price for *A. mongholicus* was predicted and analyzed using the ARIMA model in SPSS (IBM, Armonk, NY, United States). It is worth noting that we first used the price data from 2017–2020 to make a price forecast for 2021, and compared the forecast results with the price data we collected to verify the accuracy of the model. Then, we forecast the price of *A. mongholicus* in 2022 based on the data from 2017–2021.

RESULTS

Study on Ecological Suitability of *Astragalus mongholicus*

The MaxEnt model prediction results were imported into ArcGIS software (George and Fred, 1971). To guide local governments and farmers toward more effective *A. mongholicus* planting strategies in the current environment, and develop medicinal value according to the regional results of the ecological suitability of *A. mongholicus*, Guyang County administrative data were covered in ArcGIS. In calculating the model results, the influence of each ecological factor on the distribution of *A. mongholicus* was determined by analyzing the response curves of those ecological factors that had a notable contribution rate and were ranked high in importance.

Accuracy of MaxEnt Model

The ROC curve showed that the area under the curve (AUC) value of the *A. mongholicus* test sample was 0.872. According to the AUC evaluation standard, the MaxEnt model results were ideal and reached a good level, which also demonstrated the validity of the model for evaluating the habitat suitability of *A. mongholicus* (**Figure 3**).

Regional Ecological Suitability for *Astragalus mongholicus* in Guyang County

The model analysis results demonstrated that the north of Xidoupu Town, the west of Yinhao Town, and the middle of Xiashihao Town had higher ecological adaptability. Xingshunxi Town and Huaishuo Town had a large area of medium growth area suitability (**Figure 4**).

Main Ecological Factors Affecting the Growth of *Astragalus mongholicus*

Precipitation of the warmest quarter (26.6%), minimum temperature of the coldest month (20.3%), standard deviation of seasonal temperature changes (11%), range of mean annual temperature (10.8%), and mean diurnal range [mean of monthly (max temp - min temp)] (7.4%) were identified as the five environmental variables with the highest contribution rates to the MaxEnt modelling results. Moreover, their cumulative contribution was 76.1%. These results indicated that the identified environmental variables were the main ecological factors affecting the habitat of *A. mongholicus* (**Table 2**).

Content of Index Components and Relationships With Main Ecological Factors

The astragaloside IV and calycaryin-7-glucoside contents of the 37 *A. mongholicus* samples that were evaluated, are shown in the **Supplementary Table 1**. The relationships between the astragaloside IV and isoflavone glycoside contents of *A. mongholicus*, respectively, and the main ecological factors are represented by the following equations:

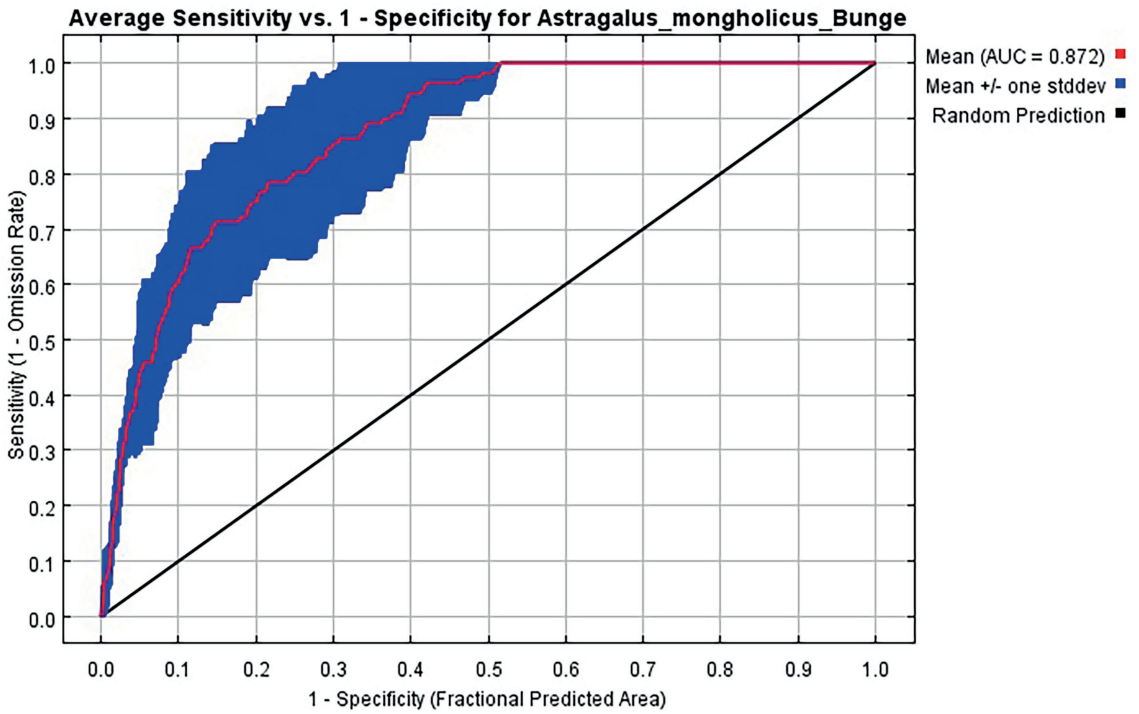


FIGURE 3 | ROC value of *A. mongholicus* modeled by MaxEnt based on distribution date.

$$y_1 = 0.024x_1 - 0.016x_2 + 0.012x_3 + 0.009x_4 + 0.007x_5 - 0.009x_6 - 0.029x_7 - 0.003x_8 - 0.003x_9 + 0.012x_{10} + 0.013x_{11} - 0.001x_{12} - 0.363$$

$$\left(\begin{array}{l} R^2 = 0.491, p \le 0.05;\ y_1 :\ calycaryin - 7 - glucoside;\ x_1 : \\ Bio1;\ x_2 :\ Bio2;\ x_3 :\ Bio3;\ x_4 :\ Bio6;\ x_5 :\ Bio7;\ x_6 : \\ Bio10;\ x_7 :\ Bio11;\ x_8 :\ Bio13;\ x_9 :\ Bio14;\ x_{10} : \\ Bio16;\ x_{11} :\ Bio18;\ x_{12} :\ Bio19 \end{array} \right).$$

$$y_2 = -0.068x_1 + 0.047x_2 - 0.005x_3 + 0.001x_4 + 0.038x_5 - 0.004x_6 + 0.001x_7 - 0.002x_8 - 0.006x_9 + 0.021x_{10} - 0.017x_{11} - 0.025x_{12} - 0.035x_{13} + 0.009x_{14} + 0.026x_{15} + 0.013x_{16} + 0.049x_{17} - 0.001x_{20} + 4.224$$

$$\left(\begin{array}{l} R^2 = 0.379,\ p \le 0.05;\ y_2 : \\ astragaloside\ \mathrm{IV};\ x_2 :\ Bio1;\ x_2 : \\ Bio2;\ x_3 :\ Bio3;\ x_4 :\ Bio4;\ x_5 : \\ Bio6;\ x_6 :\ Bio7;\ x_7 :\ Bio8;\ x_8 : \\ Bio9;\ x_9 :\ Bio10;\ x_{10} :\ Bio11;\ x_{11} : \\ Bio12;\ x_{12} :\ Bio13;\ x_{13} :\ Bio14;\ x_{14} : \\ Bio15;\ x_{15} :\ Bio16;\ x_{16} :\ Bio18;\ x_{17} : \\ Bio19;\ \mathrm{x}_{20} :\ altitude \end{array} \right).$$

These equations indicate that most ecological factors exerted different effects on the accumulation of the two active substances. For example, the mean temperature of the coldest quarter (BIO11) played an important role in the accumulation of both active compounds; it was negatively correlated with the accumulation of astragaloside IV and positively correlated with that of isoflavone glycoside. Mean annual temperature (BIO1) had a greater impact on the astragaloside IV content, to which it was negatively correlated. Mean diurnal range [mean of monthly (max temp - min temp)] (BIO2), precipitation of the coldest quarter (BIO19), and minimum temperature of the coldest month (BIO6) had positive effects on the accumulation of astragaloside IV. Moreover, mean annual temperature (BIO1) was positively correlated with the accumulation of calycosin-7-glucoside. As indicated in **Figure 5**, the most suitable areas for the accumulation of the two compounds are similar. The areas with high compound content are mainly distributed in the north and central areas of Guyang County, while the southern area is relatively small. Therefore, cultivation of *A. mongholicus* with high-quality active compounds is mainly suitable in the north of Xidoupu Town, the west of Yinhao Town, and the middle of Xiashihao Town.

Image Feature Analysis

Remote sensing technology can obtain the dynamic change information of the *A. mongholicus* planting area over time, and further provide basic data for the layout formulation of the *A. mongholicus* planting industry, and the establishment of a high-quality *A. mongholicus* cultivation technology system. Based on the spectral analysis results, months with large spectral

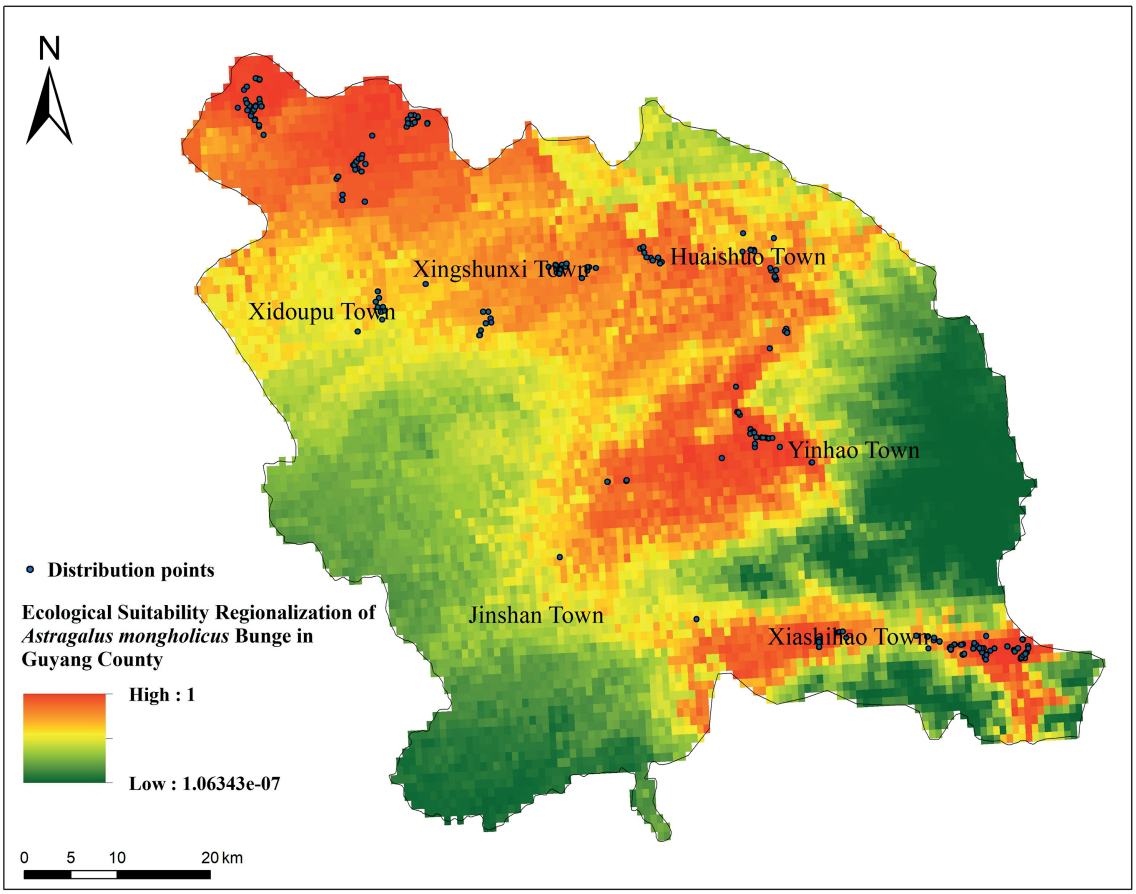


FIGURE 4 | Ecological Suitability Regionalization of *A. mongholicus* in Guyang County.

differences (June, July, August, September, and October) were further analyzed. The images were collected on June 17, July 30, August 21, September 30, and October 20, 2021. First, PCA was performed on the original images associated with the selected 5 months to obtain those bands with a larger monthly contribution rate, whereafter the texture features of these bands were analyzed. The raw images, vegetation index bands, and texture feature bands for the selected 5 months were then fused.

Principal Component Analysis

PCA was used in this study because it both aids selection of useful features and improves separability in the transformed feature space. In addition, PCA is beneficial in terms of its operation, because it is an unsupervised analysis. This method generates super pixels by simple linear iterative clustering (SLIC) and then transforms the features of super pixels using PCA. The transformed features are then used for the final classification (Su, 2019). Herein, PCA was performed on the 10 raw spectral bands for June 2021 using the "Forward PCA Rotation New static and Rotate" tool in ENVI Classic (**Table 3**). The contribution rate of the first principal component was 98.07%, and the cumulative contribution rate of the first and second principal components was 99.49%, indicating that the first two characteristics accounted for the majority. To ensure computational efficiency, feature extraction with a monthly cumulative contribution rate of 99% was selected, and used the same method to perform PCA on images from July, August, September, and October.

Texture Feature Extraction

As far as the texture features of Guyang County are concerned, the most obvious feature was the mountain, of which the coverage was extremely wide, presenting with rough texture features. In addition, the farmland was mainly distributed in relatively flat areas and in a regular state, owing to anthropogenic influences. In contrast, the grassland was distributed in a more scattered pattern and consisted mostly of herbaceous vegetation without a particularly obvious canopy structure. ENVI software was used to extract texture features from the Sentinel-2 imagery of Guyang County, and to calculate the gray level co-occurrence matrix (GLCM), thereby generating a total of eight textures: mean, variance, homogeneity, contrast, dissimilarity, entropy, angular second moment, and correlation (Xu et al., 2016). This allowed for *A. mongholicus* to be better distinguished from the other crops.

TABLE 2 | Details of the 20 ecological factors used to predict *Astragalus mongholicus* distribution.

Abbreviation	Name	Relative contribution	Type
BIO1	Mean annual temperature	0.8%	Continuous
BIO2	Mean Diurnal Range (Mean of monthly (max temp - min temp))	7.4%	Continuous
BIO3	Isothermality	0.5%	Continuous
BIO4	Standard deviation of seasonal changes in temperature	11%	Continuous
BIO5	Maximum temperature of the warmest month	0.6%	Continuous
BIO6	Minimum temperature of the coldest month	20.3%	Continuous
BIO7	Range of mean annual temperature	10.8%	Continuous
BIO8	Mean temperature of the wettest quarter	5.7%	Continuous
BIO9	Mean temperature of the driest quarter	0.9%	Continuous
BIO10	Mean temperature of the warmest quarter	0.4%	Continuous
BIO11	Mean temperature of the coldest quarter	0%	Continuous
BIO12	Mean annual precipitation	0.2%	Continuous
BIO13	Precipitation of the wettest month	2.3%	Continuous
BIO14	Precipitation of the driest month	6.1%	Continuous
BIO15	Precipitation Seasonality (Coefficient of Variation)	1.2%	Continuous
BIO16	Precipitation of the wettest quarter	1.2%	Continuous
BIO17	Precipitation of the driest quarter	0.3%	Continuous
BIO18	Precipitation of the warmest quarter	26.6%	Continuous
BIO19	Precipitation of the coldest quarter	3.2%	Continuous
BIO20	Altitude	0.5%	Continuous

Classifier Classification

The classifier was used to extract distribution information of *A. mongholicus*, and the planting area of the evaluated *A. mongholicus* was statistically identified. The sample data were randomly divided into training set (70%) and validation set (30%). The results showed that the actual planting area of Radix Astragali was very consistent with the high suitability growth area predicted by the MaxEnt, and the planting area of *A. mongholicus* in Guyang County was 29.0123 km^2. Compared with the actual planting area, the accuracy rate was 83.69%, with certain reliability (**Figure 6**).

Accuracy Evaluation

The random forest classification algorithm was used to classify the Sentinel-2 imagery time series including 143 characteristic variables (original spectral feature and vegetation index texture feature), and extracted the distribution information of *A. mongholicus* in the image. The OA of the model was 96.51%, and the Kappa accuracy was 94.96%.

Industrial Structure and VC

As a geographical symbol of agricultural products in China, *A. mongholicus*, a specialty of Guyang County, offers higher productivity and quality. Owing to its long cultivation and supply history, we found that *A. mongholicus* production practices consist of 10 mature VCs, which can be distinguished by their various composite patterns of stakeholders. Typically, *A. mongholicus* undergoes six production stages before it reaches the wholesale and retail herbal markets, as shown in **Figure 7**.

After conducting the survey, we constructed the VCs, as shown in **Figure 8**. VCs 1–3 started with independent farmers tending to their own relatively small areas of land to grow *A. mongholicus*. This represents a traditional, small-scale farmer's economic form, and is an important part of *A. mongholicus* production. In these VCs, farmers usually sell *A. mongholicus* through large suppliers (e.g., middlemen, cooperatives, and processing companies), although the yield of *A. mongholicus* products varies from year to year, which incurs high transaction and switching costs, limiting farmers' income. However, when farmers are located close to medicinal material markets or processing companies, they can sell their products directly and increase their income, as in VC 3. Most often, middlemen and processing companies finely process and grade the *A. mongholicus* purchased from farmers, subsequently selling it to pharmaceutical factories or Chinese herbal medicine markets to gain higher profits. Finally, the products are retailed to consumers through hospitals, pharmacies, and other avenues. This pattern is most common in Guyang County (Chen et al., 2021).

VCs 4–6 represented agricultural cooperatives with relatively large areas of land, which are usually composed of several farmers who manage large tracts of land through leases or other means, with the capacity to purchase more machinery and equipment than small-scale farmers. These farmers receive more financial, technical, and equipment support to grow *A. mongholicus* individually or cooperatively. Many cooperatives involved in *A. mongholicus* planting tend to sell fresh roots to middlemen or processing companies, thus achieving rapid capital recovery in the second year of planting. In addition, some cooperatives have complete facilities to manage the wholesale, and processing of raw materials. After being subjected to a series of operations, *A. mongholicus* medicinal products in these VCs are obtained and sold directly to the Chinese herbal medicine market and pharmaceutical factories, in the hopes of reducing the consumption cost of middlemen and others, thereby increasing revenue (Jiang et al., 2021).

VC 7 started with a planting company. Planting companies generally cultivate *A. mongholicus* in large plantations using a high degree of mechanization, comprehensive sprinkler irrigation systems, and warehouses storages. After the planting company

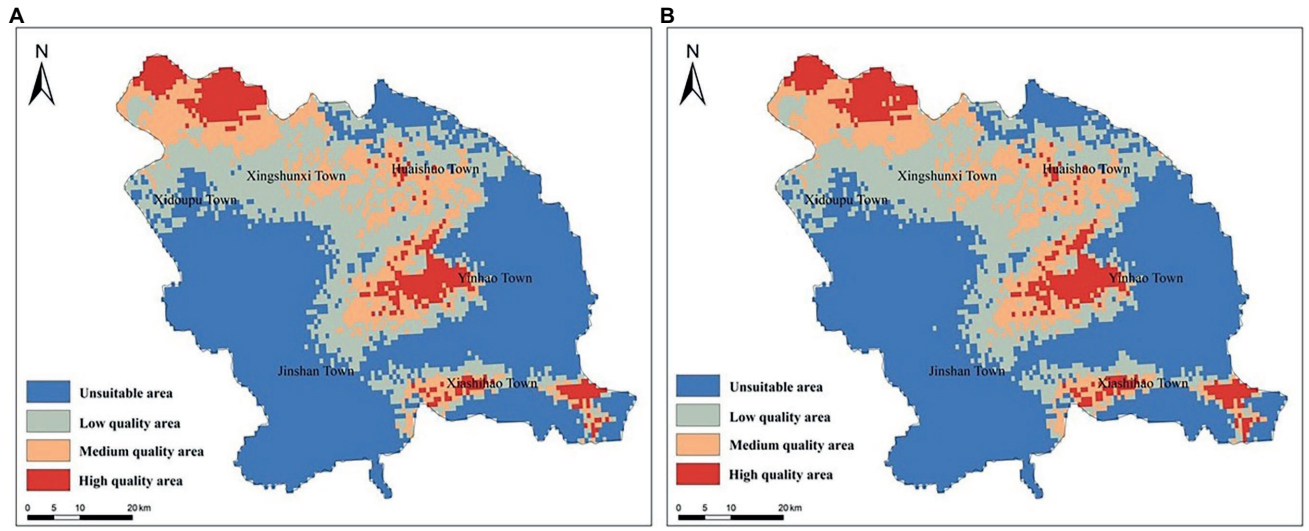


FIGURE 5 | Regional suitability distribution of high-quality *A. mongholicus* in Guyang County. Its suitability was based on the content of astragaloside IV **(A)** and calycosin-7-glucoside **(B)**.

buys the seeds, it hires farm workers to perform planting, weeding, and pest control. This pattern is more common in the area surrounding Guyang County, where large stretches of land are available for growing *A. mongholicus*.

Cultivation companies in VC 8 were involved in all VC nodes, from production and processing to wholesale. They typically rent the land and hire workers to farm, harvest, pre-process, and then process the *A. mongholicus* in the Companies' own processing plants. Processed *A. mongholicus* can be sold to hospital pharmacies and private clinics. Because these planting companies simultaneously fulfill the roles of independent farmers, middlemen, and processing entities, VCs involving them are simplified, with reduced costs and improved economic efficiency. Compared with independent farmers and agricultural cooperatives, planting companies have a more standardized planting and processing model. Moreover, all process stages are traceable, including the production chain and planting technology, application of fertilizers and pesticides, and quality inspection. The wholesale and retail sectors can control the quality and supply of *A. mongholicus* products more directly and achieve rational resource allocation and revenue maximization. In addition, most planting companies employ local, experienced farmers for planting and processing, providing job opportunities in the region. These farmers can also find other jobs during the off-season, thereby increasing their income. Planting companies around Guyang County, based on the principles of high-quality standards and timely delivery, sell processed *A. mongholicus* products to wholesale and retail departments, optimizing the supply model of *A. mongholicus* products (Bi et al., 2020).

As an export VC, VC 9 was similar to VC 8 but represented export of *A. mongholicus* products to foreign consumers. In this chain, exported *A. mongholicus* products are subject to stricter quality control measures to improve reliability of the *A. mongholicus* production system and achieve reputation and marketing goals.

VC 10 was an e-commerce-based supply model. With the development of the big data era, the internet-based Chinese medicine trade (e-commerce) is becoming increasingly common, and the future trend entails a shift of the trade center from the market to the place of origin. Contractors (in the form of middlemen or small medical processing plants) are the main players that buy pre-processed *A. mongholicus* directly from independent farmers or agricultural cooperatives, process it into medicinal slices or powder, package it more attractively, and sell it online. In theory, e-commerce should offer higher profits than traditional marketplaces because of significantly lower operating costs; however, the quality control of products sold through online platforms requires significant improvement. Moreover, e-commerce currently represents a relatively new supply model in which sales channels are not yet fully established and sales levels are far lower than those of offline sales. Further in-depth research is required to address these issues.

Price Volatility and Forecast

In recent years, the price of Chinese herbal medicines has significantly varied at high frequency, resulting in strong uncertainty and intensified market risk. The price data of Mongolian *A. mongholicus* was collected for 5 years and it was evident that the price was not fixed. Price volatility affects all stakeholders in the VCs and has a significant impact on the income for growers. Uncertainty about prices may affect growers' enthusiasm, resulting in changes in acreage. For wholesalers, price uncertainty can lead to backlogs or reduced revenues. We hope to provide stakeholders with a reference value by predicting future price trends.

TABLE 3 | Principal component analysis results of original bands in each month.

PC	Eigenvalue	Percent (%)
(a) June		
1	**********	**98.07**
2	198653.1435	**99.49**
3	50204.4439	99.85
4	7833.7430	99.91
5	5591.6879	99.95
6	3837.2299	99.98
7	1600.1198	99.99
8	823.2481	99.99
9	434.7474	100.00
10	402.0703	100.00
(b) July		
1	**********	**92.96**
2	925510.1174	**99.40**
3	54794.9855	99.78
4	10584.4943	99.86
5	7311.4799	99.91
6	6335.4789	99.95
7	3513.3720	99.98
8	1921.6643	99.99
9	945.7515	100.00
10	577.5016	100.00
(c) August		
1	**********	**93.61**
2	667177.0382	**99.41**
3	45269.5504	99.78
4	9327.8970	99.86
5	4539.5440	99.91
6	3432.8620	99.95
7	2540.6536	99.98
8	1691.1817	99.99
9	761.6588	100.00
10	352.9345	100.00
(d) September		
1	**********	**97.76**
2	178696.0256	**99.35**
3	52125.5981	99.81
4	765.1415	99.88
5	5518.4661	99.93
6	4760.6366	99.97
7	1433.4957	99.98
8	1133.7566	99.99
9	559.6960	100.00
10	487.0376	100.00
(e) October		
1	**********	**98.52**
2	102878.8513	**99.36**
3	46063.2139	99.73
4	16295.3402	99.86
5	7186.4223	99.92
6	4443.8691	99.96
7	2203.5974	99.98
8	1118.5134	99.99
9	1064.3939	100.00
10	598.6472	100.00

***** represents Eigenvalue, however, its data digits exceed the range displayed by ENVI, so the software output is ****.*

First, we predict the price in 2021 based on the ARIMA model, and get the price of *A. mongholicus* as shown in **Figure 9A**. The prices of *A. mongholicus* we collected in 2021 are all 15 yuan/kg, and the difference from the predicted value is within the range of 2 yuan/kg, which has a certain accuracy. Meanwhile, the model got a Mean Absolute Percentage Error (MAPE) of 3.369%, and R^2 was 0.948. Studies have shown that when MAPE is less than 10%, the model fit is better (Nakashima et al., 2021), so the ARIMA price prediction model is feasible.

We predicted the price of *A. mongholicus* in 2022 based on the price from 2017 to 2021, and MAPE was 2.755%, R^2 was 0.947. **Figure 9B** shows data collection date in January 2017 (horizontal axis), price of *A. mongholicus* (vertical axis), price forecast trend (blue line), and price forecast confidence interval (dotted line) is depicted. Those prices indicated after January 2022 are forecast prices. During the period from 2017 to September 2018, the price of *A. mongholicus* was relatively stable. However, during September 2018–May 2020, the price of *A. mongholicus* experienced a rapid decline and recovery. After May 2020, the price has remained stable at 15 yuan/kg. In 2022, the price of *A. mongholicus* has a gentle rise and gradually stabilized.

According to the ARIMA price forecast model, the market price of *A. mongholicus* in Mongolia may be expected to show a steady future trend, although the overall impact may not be significant. This can be gradually improved by adjusting the supply relationship related to *A. mongholicus* production, processing, and sales, thereby alleviating drastic price fluctuations.

DISCUSSION

Practical Application of SDM Prediction

The quality of medicinal plant products can be more clearly identified in different markets by understanding the internal and external linkages between production, processing, and trade networks. The added value offered by medicinal plants can be introduced at various stages of plant production. As the first step in its VC, *A. mongholicus* cultivation techniques are essential to increasing its value. In our study, the ecological and quality suitability of *A. mongholicus* was partitioned, and the quality of this medicinal plant was closely related to the choice of cultivation site. Therefore, in the process of converting medicinal plant raw materials into high-value products, the realization of value depends not only on actual production, but also on scientific guidance, which is an indispensable step (Yang et al., 2020a).

A high overlap was found between the suitable distribution area of *A. mongholicus* and the area with the highest active components. However, the environmental data that has the greatest impact on the suitable growth of *A. mongholicus* was the precipitation in the warmest quarter, although it was not the most important environmental data to promote the accumulation of its effective components. In short, although there were many similarities between the suitable distribution area and the high-quality distribution area of *A. mongholicus*, the ecological and environmental factors affecting the two were different. Studies have shown that the promotion of secondary metabolite accumulation in medicinal plant tissues is related to the interactions of multiple environmental factors (Jiang et al., 2020). The quality of many authentic medicinal plants

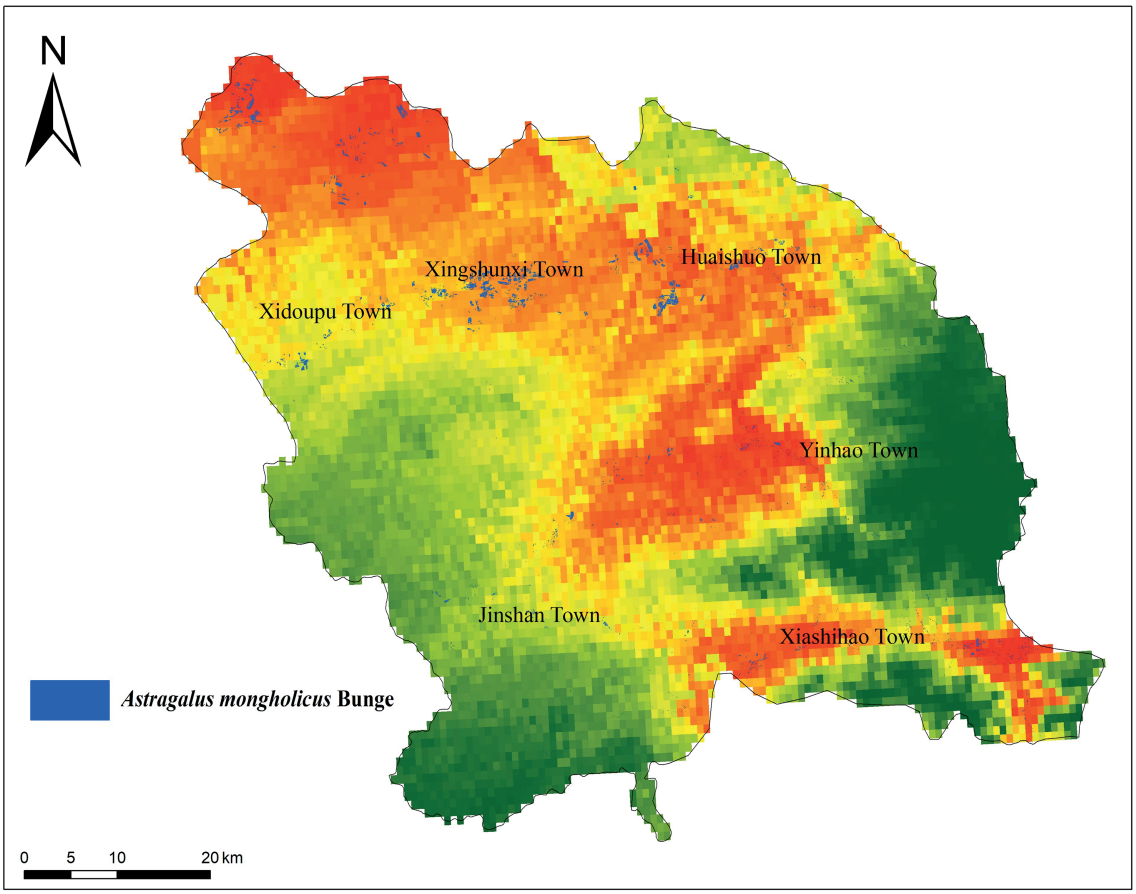


FIGURE 6 | Comparison of random forest prediction results and ecological suitability distribution results.

is related to certain environmental stressors, and it is proposed that the formation of high-quality medicinal materials may need to experience unfavorable environmental conditions (Huang and Guo, 2007) Therefore, further research is needed to explore the influence of the environment on the distribution of high-quality medicinal plants. However, the contribution of altitude to the study was not significant. It may be that most of *A. mongholicus* in Guyang County was planted in flat and broad agricultural land. The altitude range of *A. mongholicus* planting in the whole county was similar, thus the altitude had little effect. Whether altitude has an impact on the growth of *A. mongholicus* needs further study in an area with large altitude differences.

In this study, the predicted distribution of SDMs was similar to the distribution extracted from remote sensing, reflecting the major potential of using geographic information to predict the distribution of *A. mongholicus*. Furthermore, the feasibility of employing SDMs to guide the introduction and cultivation of medicinal plants was demonstrated. In a previous study, Wang et al. (2021b) used MaxEnt to estimate the impact of climate change on the distribution of potatoes in China from 1961 to 2017, predicting its suitable planting areas and guiding its planting distribution (Wang et al., 2021b).

Practical Application of Remote Sensing Prediction

Based on the Sentinel-2 imagery, the planting distribution information of *A. mongholicus* in Guyang County was extracted and the distribution and planting area of *A. mongholicus* was obtained. This method had high feasibility and can aid government departments and farmers to make informed production decisions. In particular, government departments can identify discrepancies by comparing remotely sensed planting data to those reported by farmers or businesses. In a previous study, we used ZY-3 satellite imagery as a remote sensing data source to interpret the Saposhnikovia planting situation in Naiman Banner. Thereafter, 20 sample plots were randomly selected for field measurement, and the accuracy was evaluated by comparisons with the interpretation results. Subsequent calculations revealed the accuracy of extracting the Saposhnikovia planting area at 93.90%, which meets the requirements of remote sensing monitoring of agricultural conditions (Jia et al., 2019). Remote sensing technology is widely used in agriculture; however, in the investigation of medicinal plant resources, cultivated medicinal plants are mainly monitored, whereas wild medicinal plants are rarely monitored, and a lot of experimental

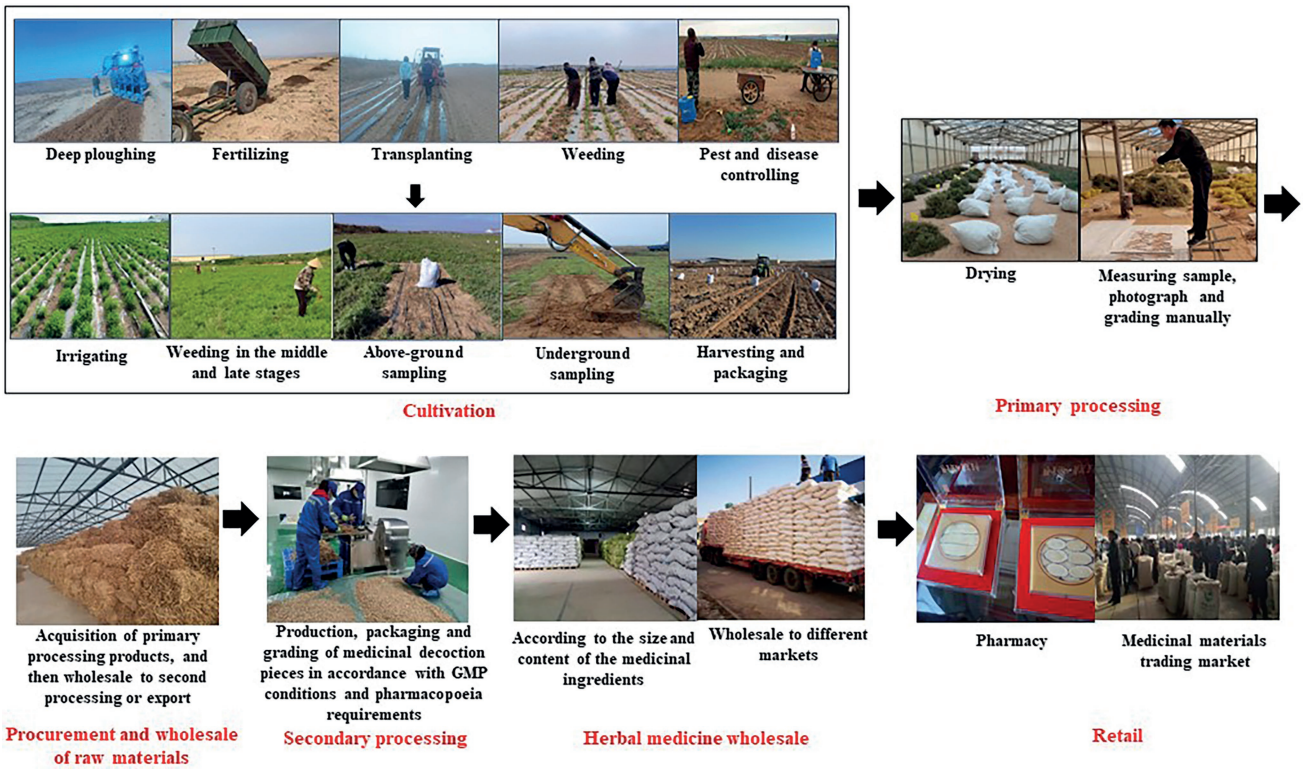


FIGURE 7 | Six stages of *A. mongholicus* before reaching consumers.

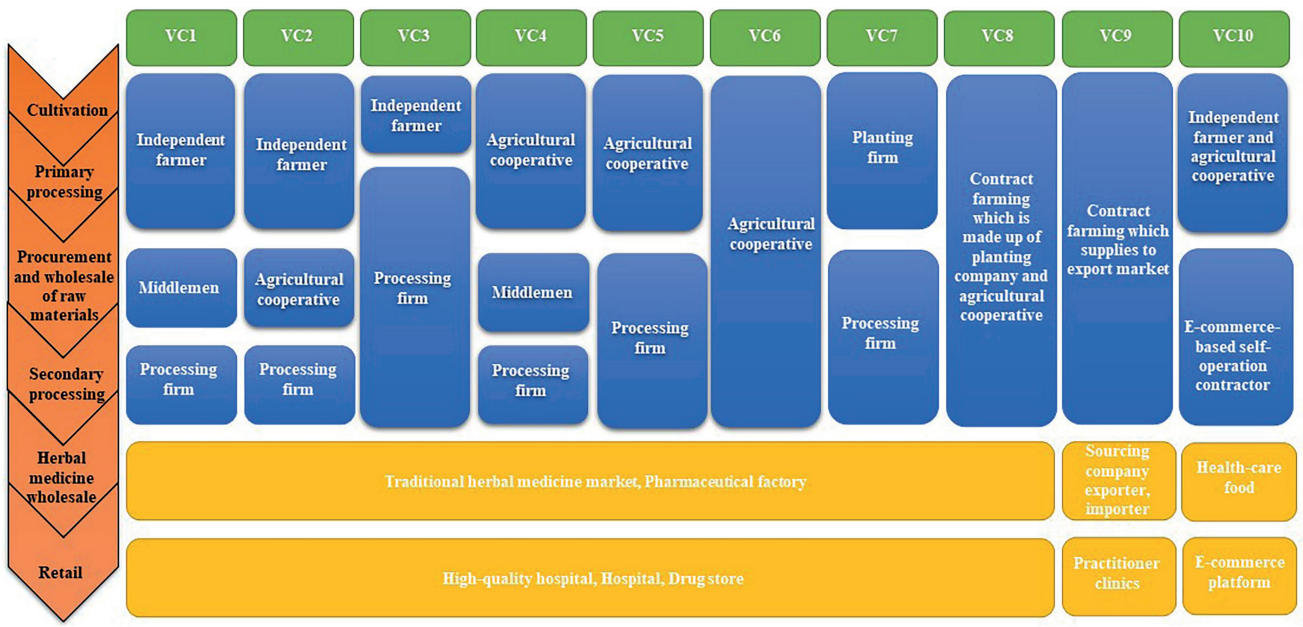


FIGURE 8 | Primary VCs and stakeholders involved in *A. mongholicus* produce.

verifications thus remain necessary. Subsequent studies should further monitor the yield, diseases, and insect pests of medicinal plants, so as to provide more statistical data for relevant management departments and promote the formulation of relevant policies and stable economic development.

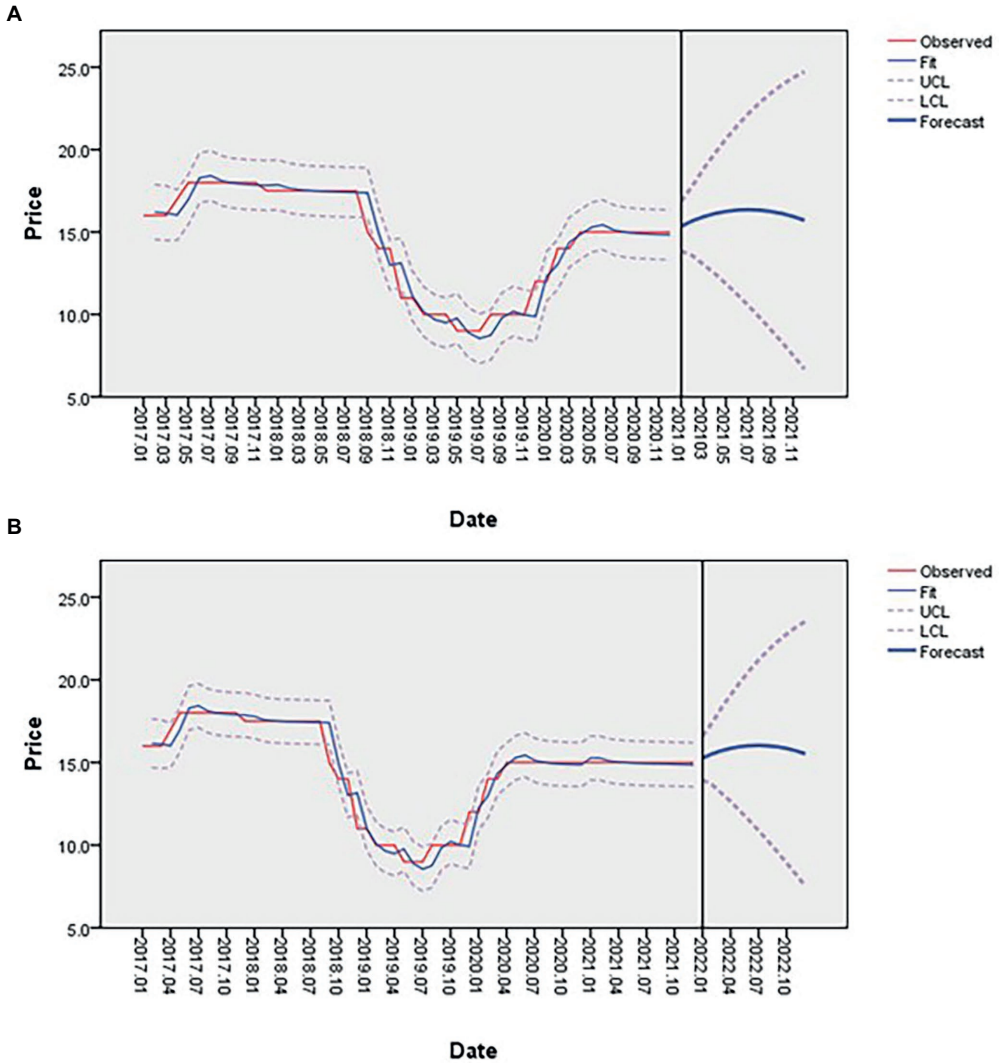


FIGURE 9 | Price forecast for 2021 **(A)**; Price forecast for 2022 **(B)**.

The Relationship Between Behavior, Income, and Quality

In traditional markets, VC stakeholders are inclined to take steps toward cost reduction and profit growth, however, this can lead to reduced product quality. VC 1 and 2, constructed herein, were typical examples of traditional markets. To achieve high yield in such VCs, excessive chemical drugs may have been used without professional guidance, and unreasonable planting measures may have been implemented, resulting in reduced quality and safety of *A. mongholicus*. Therefore, quality issues were common in these ventures, although the income of farmers was relatively low.

VCs 3–6 reflected relatively standardized market models involving larger acreages and more standardized planting patterns than that of independent farmers. However, they were highly variable and determined the quality of a medicinal materials based on the market environment.

In high-quality markets, companies were more comprehensively involved with some form of self-regulation. Consequently, their brand reputations were improved, thereby increasing their overall value. The production in VCs 7 and 8 reflected a high-quality product that would enter a high-quality market. Quality products are key to the development of a factory, and the reliability and traceability of its products allows it to build a good reputation.

In the export market, relevant companies strictly controlled every step of the production process and extensive quality tracking of *A. mongholicus* was performed by conducting germplasm selection and soil testing in the planting area. As such, it was ensured that all aspects (from planting to product packaging), met the inspection standards required by foreign markets. VC 9 implemented effective quality control measures during cultivation and production, resulting in the production of high-quality *A. mongholicus*. This would result in highly lucrative returns for the stakeholders. Selling medicinal materials

through e-commerce platforms, as in VC 10, could greatly reduce store rental costs. However, several gaps remained in the supervision process; high-quality medicinal materials from the same source could be classified into different grades and enter the market through different channels. Therefore, the behavior and interests of stakeholders, as well as the quality and target market of the product, were closely related.

The Chinese herbal medicine industry has a long chain with a wide range and many limitations. It is necessary to consider the information associated with its cultivation, processing, acquisition, storage, transportation, and sales in addition to paying more attention to the drug properties of Chinese herbal medicines. Moreover, guidance practices on the development of the Chinese herbal medicine industry should be strengthened, the large fluctuation of total output and prices caused by production dispersion should be alleviated, and the establishment of price formation mechanisms oriented by high quality and high prices should be promoted. The Chinese herbal medicine industry is highly focused on authenticity. Unlike most crops, which focus on production yield, Chinese herbal medicines are more oriented toward quality assurance. Therefore, the production of Chinese herbal medicines required unique developmental directions and ideas.

Price Fluctuation Factors

Along with improvement on the economic level, people pay increasingly more attention to their own health care and disease prevention. Market price fluctuations of traditional Chinese medicines are not only closely related to its production costs, but are also affected by factors such as policies, climate, and epidemics.

Traditional Chinese medicine plays an active role in the early interventional treatment of diseases. Moreover, according to existing literature and clinical experience, Chinese herbal medicine has a therapeutic effect on COVID-19 (Leung et al., 2020). These products have become an important part of China's fight against the pandemic, resulting in an increased demand for Chinese herbal medicine. Under the current state of prevention and control of viral dissemination, it may be difficult to harvest Chinese herbal medicines and its transportation may be inconvenienced in some areas, resulting in price fluctuations. In addition, due to the impact of bad weather patterns, drought and flood disasters have reduced production in some areas, and the worldwide spread of the epidemic has also led to insufficient supply of imports and exports (Wu and He, 2021). Overall, the demand for Chinese herbal medicines has increased; however, supply has decreased, which inevitably leads to price fluctuations. Nevertheless, it is worth noting that the planting area of Chinese herbal medicines has annually increased in recent years, which would likely improve the supply and demand problem to a certain extent and alleviate the sharp price fluctuations.

Limitations of This Study

In this study, we selected Guyang County as the study area, considering that the *A. mongholicus* grown in this area has been recognized as an agricultural geographical indication product by the Ministry of Agriculture and Rural Affairs of the People's Republic of China, and the high quality of *A. mongholicus* grown in this area has been recognized by the state and society. However, Guyang County is small, the *A. mongholicus* cultivation area is limited, and the relationship between supply and demand cannot be improved to a large extent. In order to effectively explore the suitable planting areas of *A. mongholicus* in the country, we executed a detailed investigation of various data. According to the results of the fourth census of traditional Chinese medicine resources, there exists a large number of *A. mongholicus* in Northeast, North, and Northwest China, especially in Shanxi and Gansu. Nevertheless, this study provides a reference for future research.

CONCLUSION

With increasing demand for traditional Chinese medicine in China, the sustainable development of the traditional Chinese medicine industry has received more attention. In this study, we combined remote sensing technology with SDM to predict the suitable growth area of *A. mongholicus* in Guyang County, aiming to provide theoretical guidance for the selection of *A. mongholicus* planting areas. The relationship model between the active components and ecological factors of *A. mongholicus* in Guyang County was established to evaluate the main factors affecting the accumulation of active components in *A. mongholicus*. Among these ecological factors, mean temperature of the coldest quarter (BIO11) played an important role in the accumulation of astragaloside IV. Moreover, mean diurnal range [mean of monthly (max temp - min temp)] (BIO2) and precipitation of the coldest quarter (BIO19) influenced the accumulation of calycaryin-7-glucoside. The northern and central regions of Guyang County were predicted to be suitable planting areas for high-quality *A. mongholicus*. Sentinel-2 imagery were used to monitor the growth area of *A. mongholicus*, which resembled the actual planting situation and provided data references which may be used by the relevant management departments to formulate the required policies and plans for conducting economic management. Moreover, the behavior of stakeholders, suitability of geographic planting areas, and target markets had significant influences on the quality of *A. mongholicus*. Stakeholders in each VC played different roles in the cultivation, processing, and sales of *A. mongholicus*. However, in addition to the selected planting area and associated quality, yield, pests, and diseases also affected the supply of *A. mongholicus*. In future research, remote sensing technology should be used to monitor the yield, diseases, and insect pests of *A. mongholicus*, balance supply and demand, and aid in the development of the *A. mongholicus* industry. Our research proved that it is necessary to conduct a larger-scale and more comprehensive study on the suitable cultivation areas of *A. mongholicus* by means of remote sensing and SDM, to lay a foundation for deepening its role in the environment and to allow for its sustainable utilization.

With the rise of the trend of advocating traditional Chinese medicine and natural medicine in the world, the international recognition of medicines and health food produced from natural Chinese herbal medicine resources has been increasing, and the development and utilization of Chinese herbal medicine resources has become an important trend in the development of medicine in the world. As a health care product, *A. mongholicus* can be used both as medicine and as food, and has a wide range of uses. In order to give full play to the role of *A. mongholicus*, in addition to continuing to strengthen planting techniques, we should also develop *A. mongholicus* products that meet different needs according to the uniqueness of *A. mongholicus*. Exploiting its market potential is of great significance to promoting the development of *A. mongholicus* industry, promoting regional development, and building a green *A. mongholicus* production base.

DATA AVAILABILITY STATEMENT

The datasets presented in this study can be found in online repositories. The names of the repository/repositories and accession number(s) can be found in the article/**Supplementary Material**.

AUTHOR CONTRIBUTIONS

RZ conducted the experiment and analyzed the data. MZ, YC, YY, LJ, and XW conducted the experiment. XZ, HL, and ML designed the experiment. All authors have reviewed the final version and approved it for publication.

FUNDING

This work was supported by Ministry of Agriculture and Rural Affairs of the People's Republic of China (Grant number: CARS-21); National Administration of Traditional Chinese medicine [Grant number: Finance Society (2019), 39]; National Natural Science Foundation of China (Grant number: M1942003); Inner Mongolia Mongolian Medicine Standardization Project [2020-(MB015)]; and Science and technology program in Inner Mongolia (Grant number: 2020GG0144).

ACKNOWLEDGMENTS

I would like to thank MZ and LJ for their suggestions on revision during this period. At the same time, I would like to thank Editage for helping with the English language.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fpls.2022.908114/full#supplementary-material

REFERENCES

Alabdulrazzaq, H., Alenezi, M. N., Rawajfih, Y., Alghannam, B. A., Al-Hassan, A. A., and Al-Anzi, F. S. (2021). On the accuracy of ARIMA based prediction of COVID-19 spread. *Results Phys.* 27:104509. doi: 10.1016/j.rinp.2021.104509

Bi, Y. Q., Bao, H. Y., Zhang, C. H., Yao, R. Y., and Li, M. H. (2020). Quality control of Radix Astragali (The root of Astragalus membranaceus var. mongholicus) Along its value chains. *Front. Pharmacol.* 11:562376. doi: 10.3389/FPHAR.2020.562376

Booker, A., Johnston, D., and Heinrich, M. (2012). Value chains of herbal medicines Rresearch needs and key challenges in the context of Ethnopharmacology. *J. Ethnopharmacol.* 140, 624–633. doi: 10.1016/j.jep.2012.01.039

Booker, A., Johnston, D., and Heinrich, M. (2015). Value chains of herbal medicines Ethnopharmacological and analytical challenges in a globalizing world. *Evi. Validation Herb. Med.* 29–44. doi: 10.1016/B978-0-12-800874-4.00002-7

Breiman, L. (2001). Random forests. *Mach. Learn.* 45, 5–32. doi: 10.1023/A:1010933404324

British Pharmacopoeia Commission (2019). *British Pharmacopoeia (Herbal Drug Preparations and Herbal Medicinal Products)*. England: The Stationery Office.

Chen, X. (2021). Application of satellite remote sensing Technology in Forestry Investigation and Planning. *Modern Agri. Res.* 27, 79–80. doi: 10.19704/j.cnki.xdnyyj.2021.08.037

Chen, Y., Lei, L. J., Bi, Y. Q., Jiang, L. L., Guo, W. F., Wang, J. H., et al. (2021). Quality control of Glehniae Radix, the root of *Glehnia Littoralis* Fr. Schmidt ex Miq., Along its Value Chains. *Front. Pharmacol.* 12:729554. doi: 10.3389/FPHAR.2021.729554

Cui, X. S., Ding, Y. H., Gao, X. Q., Jin, P. B., Guo, Y. H., and Dong, X. H. (2020). Empirical study on market Price classification and prediction of traditional Chinese medicine based on multiple linear regression model. *Northern Hortic.* 14, 157–161. doi: 10.11937/bfyy.20193701

Durazzo, A., Nazhand, A., Lucarini, M., Silva, A. M., Souto, S. B., Guerra, F., et al. (2021). Astragalus (astragalus membranaceus bunge): botanical, geographical, and historical aspects to pharmaceutical components and beneficial role. *Rendiconti Lincei Sci. Fisiche Nat.* 32, 625–642. doi: 10.1007/s12210-021-01003-2

European Pharmacopoeia Commission (2017). *European Pharmacopoeia*. Strasbourg: European Directorate for the Quality of Medicines and Health Care.

George, F. J., and Fred, C. C. (1971). Error on Choroplethic maps: definition, measurement, reduction. *Ann. Assoc. Am. Geogr.* 61, 217–244. doi: 10.1111/j.1467-8306.1971.tb00779.x

Hess, A. S., and Hess, J. R. (2018). Principal component analysis. *Transfusion* 58, 1580–1582. doi: 10.1111/trf.14639

Huang, L. Q., and Guo, L. P. (2007). Formation of medicinal herbs. *J. Chinese Mater. Med.* 32, 277–280. doi: 10.3321/j.issn:1001-5302.2007.04.001

Inoue, Y., and Olioso, A. (2006). Estimating the dynamics of ecosystem co2flux and biomass production in agricultural fields on the basis of synergy between process models and remotely sensed signatures. *J. Geophys. Res.* 111:D24. doi: 10.1029/2006jd007469

Jaynes, E. T. (1957). Information theory and statistical mechanics. *Phys. Rev.* 106, 620–630. doi: 10.1103/PhysRev.106.620

Jia, J. Y., Cao, R., Zhang, X. B., Shi, T. T., Yang, M., and Li, M. H. (2019). Monitoring of Saposhnikovia divaricate planting area based on texture and pop information in Naiman banner. *China J. Chinese Mater.* 44, 4111–4115. doi: 10.19540/j.cnki.cjcmm.20190731.111

Jiang, D. Q., Wang, H. Y., Kang, C. Z., Jiang, J. Y., Du, Y. X., Zhang, Y., et al. (2020). Influence and mechanism of stress combination on medicinal plants secondary metabolism. *China J. Chinese Mater. Med.* 45, 2009–2016. doi: 10.19540/j.cnki.cjcmm.20200302.106

Jiang, L. L., Zhou, B. C., Wang, X. Q., Bi, Y. Q., Guo, W. F., Wang, J. H., et al. (2021). The quality monitoring of Cistanches Herba (*Cistanche deserticola*

Ma): A value chain perspective. *Front. Pharmacol.* 12:782962. doi: 10.3389/FPHAR.2021.782962

Lan, J. X., Zhang, F., Lian, C. L., and Li, C. (2021). Application of remote sensing Technology in Medicinal Plant Resources. *Technol. Innov. App.* 11, 167–170.

Leung, E. L., Pan, H. D., Huang, Y. F., Fan, X. X., Wang, W. Y., He, F., et al. (2020). The scientific foundation of chinese herbal medicine against COVID-19. *Engineering* 6, 1099–1107. doi: 10.1016/j.eng.2020.08.009

Liu, Y. T., Li, Z. Y., Li, H. K., and Wang, X. L. (2021). A southern citrus woodland extraction method combing phenological and texture features. *Sci. Surveying Mapping* 46, 83–93. doi: 10.16251/j.cnki.1009-2307.2021.09.011

Na, R. H., Zheng, J. H., Guo, B. L., Sen, B. T., Shi, M. H., Sun, Z. Q., et al. (2013). Remote sensing estimation of safflower planting area based on PCA and texture features. *J. Chinese Mater.* 38, 3681–3686. doi: 10.4268/cjcmm20132116

Nakashima, T., Ogata, S., Noguchi, T., Tahara, Y., Onozuka, D., Kato, S., et al. (2021). Machine learning model for predicting out-of-hospital cardiac arrests using meteorological and chronological data. *Heart* 107, 1084–1091. doi: 10.1136/heartjnl-2020-318726

National Pharmacopoeia Committee (2020). *Committee for the Pharmacopoeia of PR China. Pharmacopoeia of PR China (Part I)*. Beijing: People's Health Publishing.

Ngarega, B. K., Masocha, V. F., and Harald, S. (2021). Forecasting the effects of bioclimatic characteristics and climate change on the potential distribution of *Colophospermum mopane* in southern Africa using maximum entropy (MaxEnt). *Eco. Inform.* 65:101419. doi: 10.1016/j.ecoinf.2021.101419

Phillips, S. J., Anderson, R. P., and Schapire, R. E. (2006). Maximum entropy modeling of species geographic distributions. *Ecol. Model.* 190, 231–259. doi: 10.1016/j.ecolmodel.2005.03.026

Qi, X. M. (2020). Analysis of common problems and countermeasures in planting of *Astragalus mongholicus* Bunge. *Agri. Dev. Equip.* 161–162. doi: 10.3969/j.issn.1673-9205.2020.11.077

Qin, X. M., Li, Z. Y., Sun, H. F., Zhang, L. Z., Zhou, R., Feng, Q. J., et al. (2013). Current situation and analysis of Astragalus medicinal materials resources in my country. *J. Chin. Mater.* 38, 3234–3238.

Rao, J. N. K., Box, G. E. P., and Jenkins, G. M. (1972). Time Series Analysis Forecasting and Control. *Econometrica* 40, 970–971. doi: 10.2307/1912100

Rural Social and Economic Investigation Department of National Bureau of Statistics (2021). *China Statistical Yearbook (County-level)*. Beijing: China Statistics Press.

Shi, T. T., Zhang, X. B., Guo, L. P., and Huang, L. Q. (2017). Study of extracting Peucedanum praeruptorum planted area in Ningguo of Anhui province based on multi-source and multi-phase image. *J. Chin. Mater.* 42, 4362–4367. doi: 10.19540/j.cnki.cjcmm.2017.0185

Su, T. F. (2019). Superpixel-based principal component analysis for high resolution remote sensing image classification. Multimedia tools and applications. *Multimed. Tools Appl.* 78, 34173–34191. doi: 10.1007/s11042-019-08224-6

Sun, S. Y., and Chen, G. L. (2018). Current situation, problems and countermeasures of Radix astragali industrialization in Inner Mongolia. *Mol. Plant Breeding.* 16, 5126–5133. doi: 10.13271/j.mpb.016.005126

Tan, S. X., Johnson, S., and Gu, Z. (2018). Laser depolarization ratio measurement of corn leaves from the biochar and non-biochar applied plots. *Opt. Express* 26:14295. doi: 10.1364/oe.26.014295

Vasileios, S., Ioannis, P., Charalampos, K., Alberto, A., Andrés, A. P., and Zurbano, J. A. (2018). Scalable parcel-based crop identification scheme using Sentinel-2 data time-series for the monitoring of the common agricultural policy. *Remote Sens.* 10:911. doi: 10.3390/rs10060911

Wang, C., Shi, X. Y., Liu, J. G., Zhao, J. C., Bo, X. Z., Chen, F., et al. (2021a). Interdecadal variation of potato climate suitability in China. *Agric. Ecosyst. Environ.* 310:107293. doi: 10.1016/j.agee.2020.107293

Wang, Z. J., Wei, M. S., Guo, L. F., Wang, X. Q., and Gao, L. (2021b). Extraction of crop spatial distribution information based on principal component analysis. Geomatics & Spatial Information. *Technology* 44, 114–119. doi: 10.3969/j.issn.1672-5867.2021.06.031

Wu, W. X., and He, L. Y. (2021). The general increase in the price of Chinese medicinal materials continues, and the northern flood has become a major incentive. *21st Century Bus. Herald.* 6, 1–2. doi: 10.28723/n.cnki.nsjbd.2021.004624

Xu, X., Li, J., Huang, X., Mura, M. D., and Plaza, A. (2016). Multiple morphological component analysis based decomposition for remote sensing image classification. *IEEE Trans. Geosci. Electron.* 54, 3083–3102. doi: 10.1109/TGRS.2015.2511197

Xu, Z. L., Peng, H. H., and Peng, S. Z. (2015). The development and evaluation of species distribution models. *Acta Ecol. Sin.* 35, 557–567. doi: 10.5846/stxb201304030600

Yang, M., Li, Z. Y., Liu, L. B., Bo, A., Zhang, C. H., and Li, M. H. (2020a). Ecological niche modeling of *Astragalus membranaceus* var. mongholicus medicinal plants in Inner Mongolia. *China. Sci. Rep.* 10:12482. doi: 10.1038/s41598-020-69391-3

Yang, B., Zhang, Q. J., Wang, B., Gong, X., and Zhang, Y. B. (2020b). The habitat suitability evaluation of *Salweenia bouffordiana* based on MaxEnt model. *Acta Ecol. Sin.* 40, 6077–6085. doi: 10.5846/stxb201906031167

Yao, R. Y., Heinrich, M., Wang, Z. G., and Weckerle, C. (2018). Quality control of goji (fruits of *Lycium barbarum* L. and *L. chinense* mill.): A value chain analysis perspective. *J. Ethnopharmacol.* 224, 349–358. doi: 10.1016/j.jep.2018.06.010

Zhang, M. X., Jiang, D., Yang, M., Ma, T., Ding, F. Y., Hao, M. M., et al. (2021). Influence of the environment on the distribution and quality of Gentiana dahurica Fisch. *Front. Plant Sci.* 12:706822. doi: 10.3389/fpls.2021.706822

Zhao, Y. S., Li, Z. Y., Na, M. H., Han, Q. H., Ren, K., Zhang, C. H., et al. (2017). Analysis and suggestion of ecological planting status about Chinese-Mongolian traditional medicine in Inner Mongolia. *Modern Chin. Med.* 19, 901–906. doi: 10.13313/j.issn.1673-4890.2017.7.002

Zhu, G. P., Liu, G. Q., Bu, W. J., and Gao, Y. B. (2013). Ecological niche modeling and its applications in biodiversity conservation. *Biodivers. Sci.* 21, 90–98. doi: 10.3724/SP.J.1003.2013.09106

Zhu, S. D., Peng, H. S., Guo, L. P., Xu, T. R., Zhang, Y., Chen, M. L., et al. (2017). Regionalization of Chinese material medical quality based on maximum entropy model: A case study of *Atractylodes lancea*. *Sci. Rep.* 7:42417. doi: 10.1038/srep42417

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

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

frontiers | Frontiers in Plant Science

ORIGINAL RESEARCH
published: 07 July 2022
doi: 10.3389/fpls.2022.920284

Grid Search for Lowest Root Mean Squared Error in Predicting Optimal Sensor Location in Protected Cultivation Systems

*Daniel Dooyum Uyeh[1,2,3], Olayinka Iyiola[3,4], Rammohan Mallipeddi[5], Senorpe Asem-Hiablie[6], Maryleen Amaizu[7], Yushin Ha[1,2,3] and Tusan Park[1,3]**

[1] Department of Bio-Industrial Machinery Engineering, Kyungpook National University, Daegu, South Korea, [2] Upland-Field Machinery Research Center, Kyungpook National University, Daegu, South Korea, [3] Smart Agriculture Innovation Center, Kyungpook National University, Daegu, South Korea, [4] Department of Hydro Science and Engineering, Technische Universität Dresden, Dresden, Germany, [5] Department of Artificial Intelligence, School of Electronics Engineering, Kyungpook National University, Daegu, South Korea, [6] Institutes of Energy and the Environment, The Pennsylvania State University, University Park, PA, United States, [7] College of Science and Engineering, University of Leicester, Leicester, United Kingdom

OPEN ACCESS

Edited by:
Chuanlei Zhang, Tianjin University of Science and Technology, China

Reviewed by:
Kai Huang, Nanjing Agricultural University, China
Wali Khan Mashwani, Kohat University of Science and Technology, Pakistan
Thippa Reddy Gadekallu, VIT University, India
Li Yan, Zhongyuan University of Technology, China

***Correspondence:**
Tusan Park
tusan.park@knu.ac.kr

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

Received: *14 April 2022*
Accepted: *07 June 2022*
Published: *07 July 2022*

Citation:
Uyeh DD, Iyiola O, Mallipeddi R, Asem-Hiablie S, Amaizu M, Ha Y and Park T (2022) Grid Search for Lowest Root Mean Squared Error in Predicting Optimal Sensor Location in Protected Cultivation Systems. Front. Plant Sci. 13:920284. doi: 10.3389/fpls.2022.920284

Irregular changes in the internal climates of protected cultivation systems can prevent attainment of optimal yield when the environmental conditions are not adequately monitored and controlled. Key to indoor environment monitoring and control and potentially reducing operational costs are the strategic placement of an optimal number of sensors using a robust method. A multi-objective approach based on supervised machine learning was used to determine the optimal number of sensors and installation positions in a protected cultivation system. Specifically, a gradient boosting algorithm, a form of a tree-based model, was fitted to measured (temperature and humidity) and derived conditions (dew point temperature, humidity ratio, enthalpy, and specific volume). Feature variables were forecasted in a time-series manner. Training and validation data were categorized without randomizing the observations to ensure the features remained time-dependent. Evaluations of the variations in the number and location of sensors by day, week, and month were done to observe the impact of environmental fluctuations on the optimal number and location of placement of sensors. Results showed that less than 32% of the 56 sensors considered in this study were needed to optimally monitor the protected cultivation system's internal environment with the highest occurring in May. In May, an average change of −0.041% in consecutive RMSE values ranged from the 1st sensor location (0.027°C) to the 17th sensor location (0.013°C). The derived properties better described the ambient condition of the indoor air than the directly measured, leading to a better performing machine learning model. A machine learning model was developed and proposed to determine the optimal sensors number and positions in a protected cultivation system.

Keywords: air-vapor mixture, artificial intelligence, greenhouse, machine learning, psychrometric properties, RMSE, time-series big data

INTRODUCTION

The changing climate and depletion of natural resources such as fossil-based energy, land, and water necessitate improving resource use efficiency. Protected cultivation systems such as greenhouses could be essential in efficiently providing nutritious fresh foods for a growing world population (Stanghellini, 2013). Higher water use efficiency per unit area of crop production has been recorded in protected cultivation systems compared to open-field cultivation (Li et al., 2010). This could be a potential solution to land scarcity. Where disasters such as pandemics make farms momentarily less accessible, remotely controlled and autonomous cultivation strategies would be beneficial.

However, the benefits in these systems come at higher energy demands, especially when poor decisions are made based on incorrect monitoring of the micro-climate. Overheating and consequently poor plant growth and ensuing economic losses could be one such result (Park and Park, 2011). Protected cultivation systems could, however, be capital intensive. Improved efficiency will reduce the system's energy consumption and reduce production costs (DeFacio et al., 2002; Vox et al., 2010).

In protected cultivation systems, irregular changes or high fluctuations in indoor climatic conditions can be deleterious to productivity. Temperature and relative humidity management to meet specific plant requirements is critical for survival, optimum growth, and enhanced productivity (DeFacio et al., 2002; Vox et al., 2010). The optimal placement of the minimum number of sensors for measuring the micro-climate of protected cultivation systems is critical for their efficient use and sustainability. The protected cultivation system has a high level of variability caused by plant respiration and heating systems.

Ventilation causes air movement and consequently the uniformity of the environment. In Guzmán et al. (2019), the wind direction was reported to have a significant effect on ventilation rate, airflow, and crop temperature distributions. Also, in Li et al. (2010), it was observed that temperature did not rise linearly between inlet and fans and was higher at or above the top of the crop canopy than within it in a full-size house but not in a glasshouse compartment. A method for determining the optimal number and locations of the sensors would be necessary to accurately measure the environment of a protected cultivation system.

Recent high-tech protected cultivation systems are equipped with advanced sensors for monitoring parameters such as temperature, relative humidity, CO_2, and light. This is done to improve monitoring and control of micro-climate parameters and sometimes facilitate remote-controlled and autonomous cultivation. Decisions may be made based on various actuators used to regulate heating, lighting, cooling, dosing of CO_2 and fertilizers, dehumidification, irrigation, screening, fogging, as examples (Nelson, 1991; Uyeh et al., 2019, 2021; Bhujel et al., 2020; Gadekallu et al., 2021). These actuators operate based on sensors providing feedback on measured data for the control loop set points configured in a computing device (Stanghellini, 2013; Graamans et al., 2018).

In autonomous growing systems (Stanghellini, 2013; Graamans et al., 2018; Hemming et al., 2020), deployment of the more costly, high-precision sensors have added benefits such as durability and reduced capital costs in the long-term. Decisions based on imprecise measurements could result in poor plant growth (due to under-or over-heating) or irreversible damage and associated economic losses. An additional benefit of using more precise sensors is energy savings.

Growers constantly face decision-making and optimization problems in agriculture. Multiclass models have been used to develop multivariate statistical methods in agriculture (Guzmán et al., 2019) and Principal Component Analysis - whale optimization-based neural networks to classify diseases in plants (Li et al., 2010). Others include algorithms and systems for improved decision-making and optimizations (Nelson, 1991; DeFacio et al., 2002; Vox et al., 2010; Park and Park, 2011; Uyeh et al., 2019; Gadekallu et al., 2021). Machine learning provides opportunities to solve complex tasks such as optimal sensor placement because of its capabilities to efficiently compute vast and complex datasets with a high success ratio and fewer errors (Syed and Hachem, 2019a,b).

To solve the optimal sensors placement problem, this study, (a) designed and fabricated temperature and humidity sensors to monitor every section of a protected cultivation system and accurately collect data per minute were, (b) derived psychometric properties to understand better, the actual condition and behavior of the air-vapor mixture in a protected cultivation system, and (c) proposed a machine-learning solution based on the derived psychometric properties.

A machine learning algorithm, the Gradient Boosting Algorithm, was implemented as a multi-objective approach to determine the optimal number of sensors and locate their best position. The objective function of this algorithm was to minimize the root mean squared error (RMSE) and the number of sensors using two multiple hyper-parameter tuning algorithms (Random Search and Grid Search).

Related Works

Growers constantly face decision-making and optimization problems. Multiclass models have been used to develop multivariate statistical methods in agriculture (Gadekallu et al., 2021) and principal whale optimization-based neural networks to classify diseases in plants (Gadekallu et al., 2021). Others include algorithms and systems for improved decision-making and optimizations (Park et al., 2019; Syed and Hachem, 2019a,b; Uyeh et al., 2019, 2021; Bhujel et al., 2020). Using an inadequate number of sensors may lead to under-performance, while a likely result of being superfluous is large sizes of redundant data and its associated management problems. The sensor placement problem has been recognized and studied in other fields. These include fire detection in a target region (Li et al., 2013), air and water quality monitoring (Du et al., 2014; Fontanini et al., 2016), and monitoring physical activity in humans with a three-dimensional accelerator (Boerema et al., 2014). Others include structural health monitoring based on modal data (Chang and Pakzad, 2014; Tong et al., 2014) and mid and low frequency range methods (Rao et al., 2014). Attempts have been made to

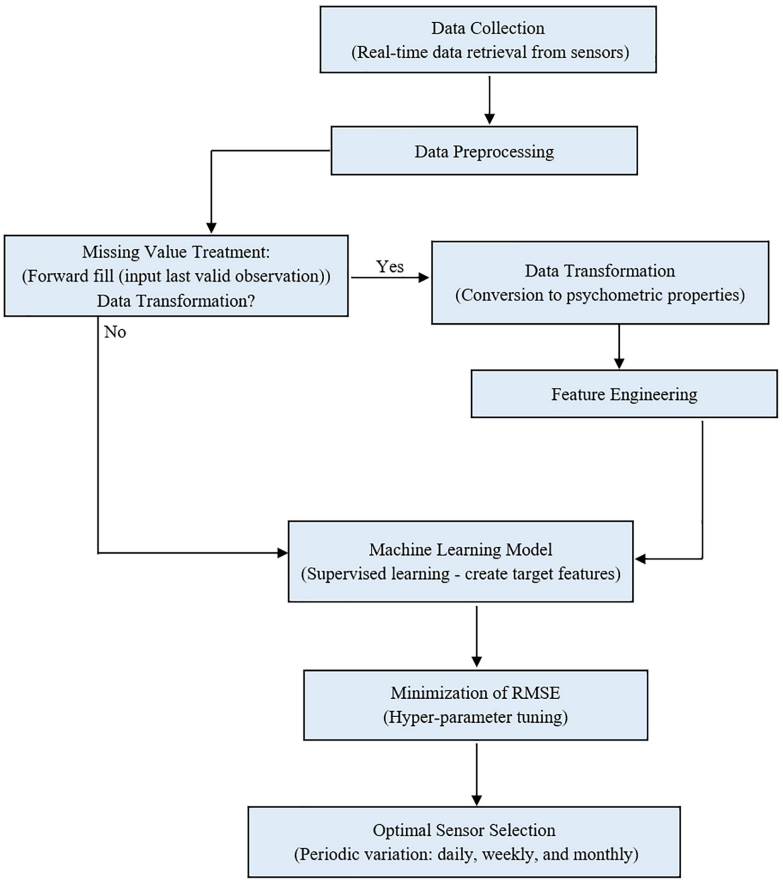


FIGURE 1 | Workflow for optimal sensor selection using (state method) in a protected cultivation system.

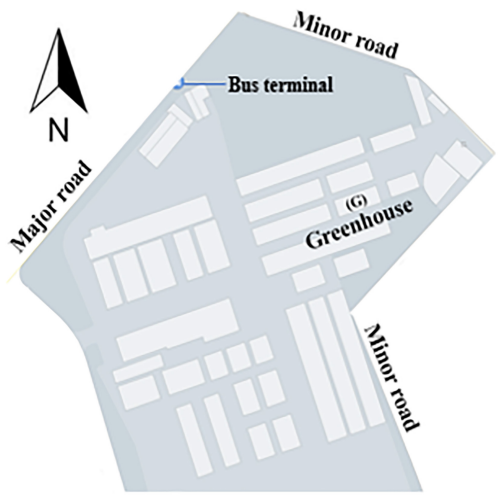


FIGURE 2 | Location of the experimental greenhouse (G) used for data collection for optimal sensor placement study.

determine the optimal sensor selection and location in internal environments, focusing on structures stability (Worden and Burrows, 2001; Löhner and Camelli, 2005; Wang et al., 2009; Chang et al., 2012; Hu and Patel, 2014; Huang et al., 2014; Arnesano et al., 2016; Seabrook, 2016). Worden and Burrows (2001) studied the optimal temperature sensor location using an error-based approach for monitoring a stadium's heating, venting, and air-conditioning systems.

As the environment in protected cultivation systems is dynamic, optimal sensor placement may involve the following scenarios: (a) multiple sensor types required in one system with two or more sometimes embedded as one (Faris and Mahmood, 2014); (b) movements of the rising and setting sun which affects the internal data (Cossu et al., 2014; Wang et al., 2014); (c) multiple layers of plant beds with varying atmospheric conditions at each level (Pamungkas et al., 2014); and (d) the influence of other internal structures of the system.

Techniques for selecting and installing sensors for monitoring and controlling climatic conditions in protected cultivation systems such as plant factories, greenhouses, etc., have been mostly heuristic. Feng et al. (2013), simulated greenhouse internal air temperature and wind-velocity distributions and suggested that the optimal sensor location is where the air and speed do not change rapidly. Several approaches, such as z-index, the outliers, and statistical measures, including central tendency and dispersion measures, have been employed (Lee et al., 2019). This study's limitation was the low volume of air temperature data and

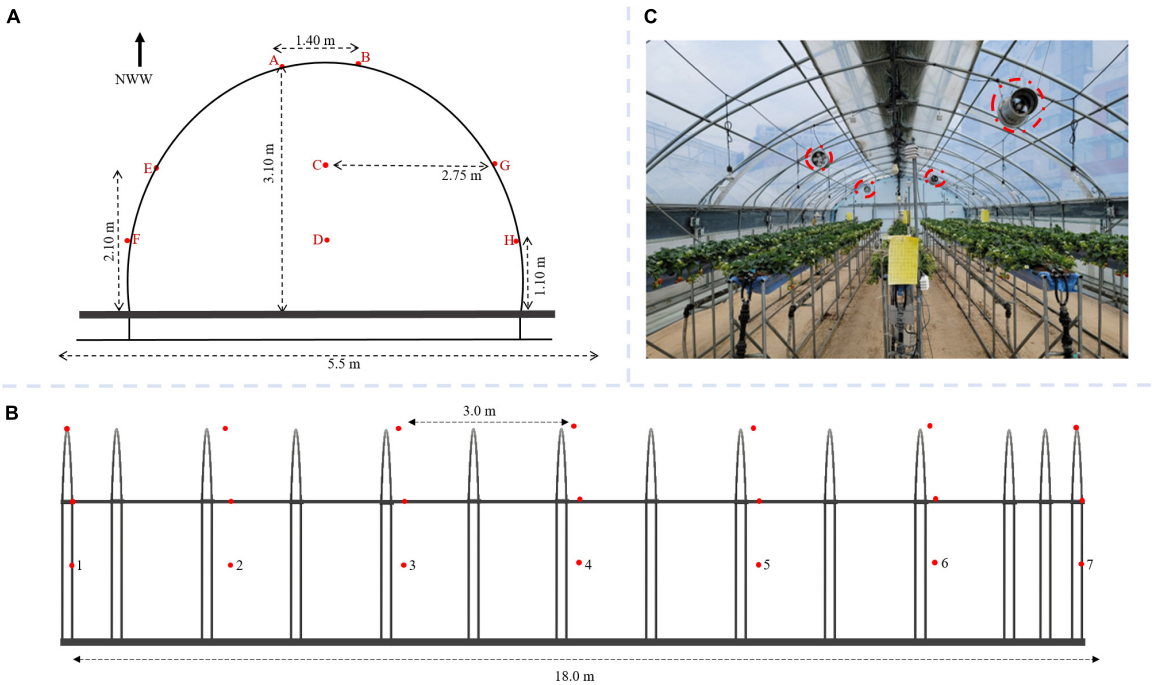


FIGURE 3 | The experimental greenhouse with temperature and relative humidity sensors installed; **(A)** front view; **(B)** side view for optimal sensor placement study; and **(C)** with growing strawberry plants and fan for mechanical ventilation circled in broken red lines.

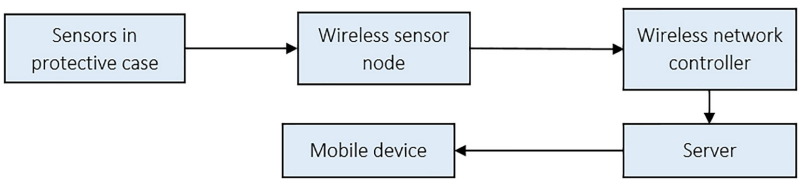


FIGURE 4 | Wireless system architecture for remote sensing of the protected cultivation system.

the non-inclusion of other influencing environmental variables such as humidity and light. In more complex and larger-sized systems, statistically based techniques incapable of handling big data would be ineffective.

Some studies attempted to use machine learning to determine the number of optimal sensors and identify their locations (Aydin et al., 2019), however, derived conditions (dew point temperature, humidity ratio, enthalpy, and specific volume), or some other environmental variables were not taken into consideration to provide a better representation of the protected cultivation system state. According to Ponce et al. (2014), most analytic models focusing on controlling the internal environment of protected cultivation systems have been based on a state-space relationship. This state-space form includes variables such as indoor temperature, humidity, energy input, outdoor temperature, wind speed, time, etc. Further, they (Ponce et al., 2014) recorded temperature and humidity are influential variables used to simplify the greenhouse state. Psychometric properties such as dew point temperature, humidity ratio, enthalpy, and specific volume would be beneficial to better represent the greenhouse's dynamic behavior, especially since air is mixed with vapor (Czubinski et al., 2013).

METHODOLOGY

a. Overview

Temperature and relative humidity data were collected remotely from a protected cultivation system located on the research farm of Kyungpook National University, South Korea. Data was collected over seven months (February, March, April, May, June, July, and October). The time-series observation for the two conditional parameters recorded per minute were representative data.

The temperature and humidity data were preprocessed, and four psychometric variables (dew point temperature, humidity ratio, enthalpy, and specific volume) were derived and used to model the protected cultivation system's indoor environment. The algorithm was trained on 70 % of the data to ensure generalization and no overfitting. The metric of evaluation,

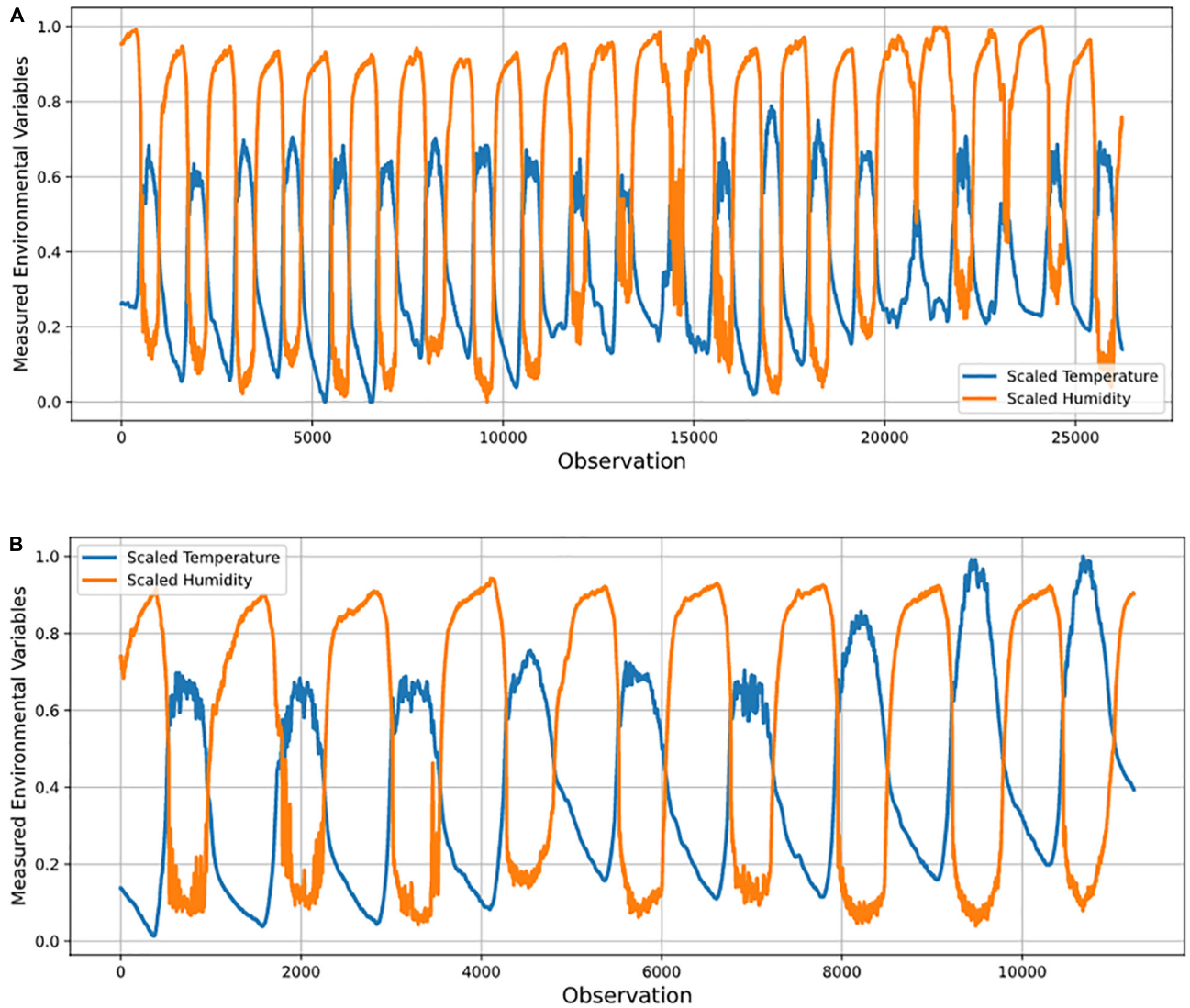


FIGURE 5 | Graphs of **(A)** training data; and **(B)** test data used in data preprocessing for sensor A1 in optimal sensor placement study.

RMSE, was minimized by tuning the algorithm's hyper-parameters (parameters whose values are used to alter the machine learning algorithm's learning rate) iteratively. Based on each month, sensor ranking was carried out. Furthermore, the number of optimal sensors required daily, weekly, and monthly was determined in a supervised manner. **Figure 1** shows the workflow for optimal sensor selection. The data is collected using the fabricated temperature and humidity sensors and stored in a cloud system. The collected data was preprocessed using forward fill and transformed into psychrometric variables. The preprocessed temperature, humidity, and transformed psychrometric data were used to develop the supervised machine learning model. Optimal sensor selection was done by minimizing RMSE using hyper-parameter tuning.

b. Experiment setup and protected cultivation system location

A Quonset-shaped protected cultivation system (greenhouse) located on the research farm of Kyungpook National University, Daegu, South Korea (35°53′43.0 N and 128°36′49.1 E) was selected for this study. The greenhouse is used to cultivate strawberries and is close to two inner roads and a major road with heavy vehicular traffic (**Figure 2**).

Fifty-six 2-in-1 temperature and humidity sensors were installed on eight rows and seven columns, each at 3 m horizontal and 1 m vertical distance apart for uniformity (**Figure 3**). The sensors were specifically manufactured to have a similar range (and error) of −20°C to 80°C (± 0.3°C) and 0% to 100% (± 2%) for temperature and relative humidity, respectively. The sensors were installed in different columns represented with A – H (**Figure 3A**) and seven fixed rows in **Figure 3C**. To prevent solar radiation from interfering with readings and causing errors, the

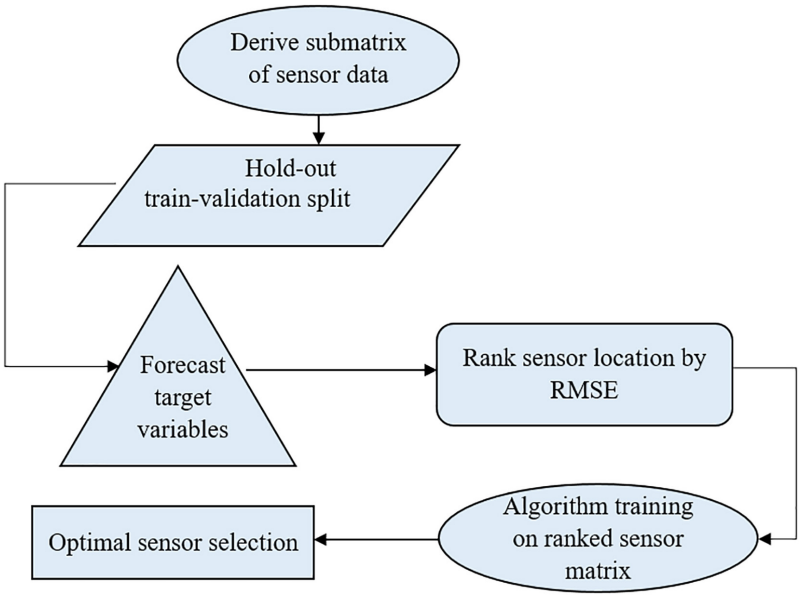


FIGURE 6 | Flow chart showing the summary of the model building process for optimal sensor selection.

sensors were enclosed in a plastic covering. Constantly running ventilation fans were installed in the greenhouse (**Figure 3B**).

c. Environmental sensing of protected cultivation system and data collection

A network controller (U-NWC-W-7S, UBN, Daegu, South Korea) was installed to minimize the temperature and relative humidity data collection error from the 56 sensors. The controller has a distributed processing system, a radio frequency of 447.9 MHz, enabling mobile software development for real-time data retrieval from the sensors. The sensor-controller system's architecture is shown in **Figure 4**.

The sensors were tightly installed to prevent movement and connected via cables to the sensor nodes, which transferred data *via* gateways to a server and then to a mobile telephone device. This wireless system enabled consistent remote monitoring. Preventive maintenance of the systems was regularly carried out to avert errors from factors such as sensor clogging.

d. Variability analysis of greenhouse environmental data

The variability of the conditions within the greenhouse was measured by calculating the Coefficient of Variation (CV) as the ratio of the standard (Equation 1) deviation to the mean temperature/humidity in each period (when expressed as a percentage) as used by Ayalew et al. (2012) and Kassie (2014).

$$Coefficient\ of\ Variation\ (CV) = \frac{Standard\ deviation, \sigma}{Mean, \mu} \quad (1)$$

e. Dynamic time warping to determine the effect of the plants on microclimate distribution

i. Data Description

Using the hourly reading of the temperature and relative humidity data collected in March with plants and June when the greenhouse was without plants, the data dimensions for March and June were 744, 113, and 720, 113, respectively.

TABLE 1 | Coefficient of Variation for temperature-relative humidity data for estimating the variability of the greenhouse.

Month	Temperature CV (%)	Relative humidity CV (%)
June	22.1	36.70
October	25.26	38.96
February	40.43	32.53
July	14.08	19.30
March	42.30	42.09
May	24.65	38.31

ii. Implementation of dynamic time warping algorithm

The dynamic time warping (DTW) algorithm, following Furlanello et al. (2006) and given below, was implemented to ascertain the effect of the plants on the microclimate distribution of the greenhouse.

Input:

series: u = $\{u_1, u_2, ..., u_{T_u}\}$

series: v = $\{v_1, v_2, ..., v_{T_v}\}$

Base conditions :

g (0,0) = 0

g (1,1) = d (u_1, v_1)·w_D

g (i,0) = ∞ for $1 \leq i \leq T_u$

g (0, j) = ∞ for $1 \leq j \leq T_v$

Recursive relation:

$$g(i,j) = \min \begin{cases} g(i, j-1) + d(u_i, v_j) \cdot w_v \\ g(i-1, j-1) + d(u_i, v_j) \cdot w_D \\ g(i-1, j) + d(u_i, v_j) \cdot w_H \end{cases}$$

for $1 \leq i \leq T_u$ and $1 \leq j \leq T_v$

Alignment deduction by tracing back from g (T_u, T_v) to g (0,0).

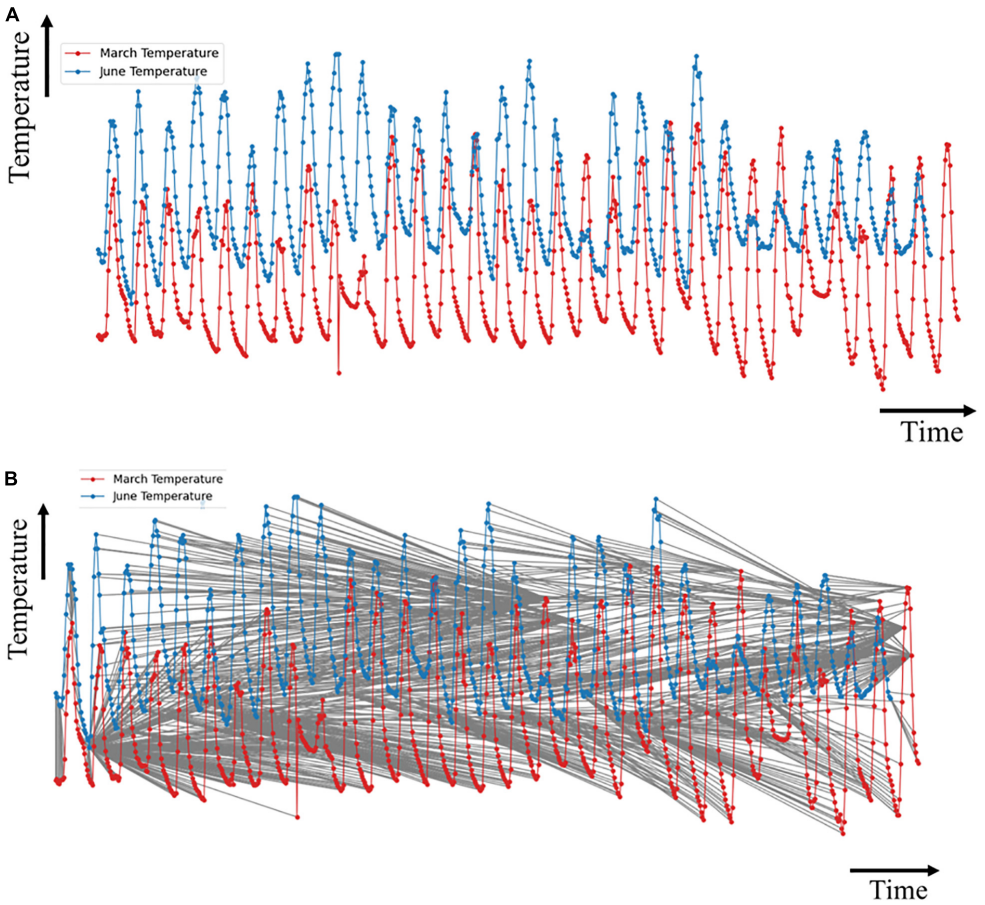


FIGURE 7 | Plots of temperature data for **(A)** March and June; and **(B)** optimal match of the time series.

Where T_u and T_v are the time points for series u and v, respectively; d is the local distance minimized by the DTW algorithm to find the minimum cost path or best alignment; g is the matrix of the dynamic table construction of $(T_u + 1) \times (T_v + 1)$; w_H, w_D, w_V are the weight configuration for horizontal (H), diagonal (D) and vertical (V) time distortions.

f. Protected Cultivation Environmental Data Preprocessing

Preprocessing the data involved standardizing features (sensor locations) and treating missing values. To standardize the features within a range of 0 to 1, feature scaling was done. Train-validation split was carried out in a time series to avoid a randomized or highly stochastic output. The tree-based algorithm (Gradient Boosting) was fitted on the training data and validated on the remaining (or unseen) portion to prevent overfitting.

Missing data was less than 1%, and these were treated with forward (or backward) filling given the appropriateness of this approach for the observations recorded within a minute. **Figure 5** shows the result of the data preprocessing at sensor A1.

From the two condition parameters – temperature and humidity – psychometric properties (dew point temperature, humid ratio, enthalpy, and specific volume) describing the air vapor mixture (Czubinski et al., 2013) in the greenhouse were derived. This helped to determine more features of importance as condition parameters for the greenhouse environment.

g. Derivation of Psychometric Variables

Equations 2–5 were used to convert the raw temperature and relative humidity data into dew point temperature, humidity ratio, enthalpy, and specific volume (Handbook, 2001):

$$\text{Dew point temperature (℃)}, \mathrm{T_d} = \frac{\mathrm{T} - (100 - \mathrm{RH})}{5} \quad (2)$$

$$\text{Humidity ratio}, \mathrm{w} = \frac{0.62198\mathrm{P_w}}{\mathrm{P} - \mathrm{P_w}} \quad (3)$$

$$\text{Enthalpy(kJ/kg)}, \mathrm{h} = 1.006\mathrm{T} + \mathrm{w}\,(2501 + 1.805\,(\mathrm{T})) \quad (4)$$

$$\text{Specific volume } (\mathrm{m^3/kg}), \mathrm{v} = \frac{\mathrm{R_{da}T}}{\mathrm{P} - \mathrm{P_w}} \quad (5)$$

Where T was internal temperature; RH, relative humidity; P_w, partial pressure of water vapor; P, total pressure; and R_{da}, gas constant for dry air = 287.055 J/(kg K).

TABLE 2 | Performance of a sensor network in identifying the optimal number of sensors and placement for measuring greenhouse conditions across different months using temperature data.

Index	Sensor location (s)	RMSE (°C)
February (F7)		
0	C2	0.0448124
1	C2, H7	0.0133850
2	C2, H7, B7	0.0114320
3	C2, H7, B7, A1	**0.0102077**
4	C2, H7, B7, A1, D6	0.0108046
5	C2, H7, B7, A1, D6, F1	0.0110145
March (G7)		
0	H7	0.0380515
1	H7, D5	0.0278228
2	H7, D5, F7	0.0239464
3	H7, D5, F7, A7	0.0215741
4	H7, D5, F7, A7, B7	0.0196743
5	H7, D5, F7, A7, B7, B4	**0.0194982**
6	H7, D5, F7, A7, B7, B4, C7	0.0199467
7	H7, D5, F7, A7, B7, B4, C7, E7	0.0200142
April (F7)		
0	D6	0.0397207
1	D6, D7	0.0349923
2	D6, D7, F6	0.0288990
3	D6, D7, F6, F4	0.0286996
4	D6, D7, F6, F4, H7	0.0285790
5	D6, D7, F6, F4, H7, G7	0.0270916
6	D6, D7, F6, F4, H7, G7, F5	0.0268920
7	D6, D7, F6, F4, H7, G7, F5, F3	0.0266292
8	D6, D7, F6, F4, H7, G7, F5, F3, E6	0.0260851
9	D6, D7, F6, F4, H7, G7, F5, F3, E6, E7	0.0226352
10	D6, D7, F6, F4, H7, G7, F5, F3, E6, E7, H6	0.0224960
11	D6, D7, F6, F4, H7, G7, F5, F3, E6, E7, H6, D2	0.0224442
12	D6, D7, F6, F4, H7, G7, F5, F3, E6, E7, H6, D2, D5	**0.0223317**
13	D6, D7, F6, F4, H7, G7, F5, F3, E6, E7, H6, D2, D5, C6	0.0226247
14	D6, D7, F6, F4, H7, G7, F5, F3, E6, E7, H6, D2, D5, C6, D3	0.0226412
May (D4)		
0	A2	0.0291603
1	A2, B2	0.0288598
2	A2, B2, B3	0.0266414
3	A2, B2, B3, F1	0.0235686
4	A2, B2, B3, F1, D3	**0.0171915**
5	A2, B2, B3, F1, D3, H1	0.0172351
6	H7, D5, F7, A7, B7, B4, C7	0.0173482
June (D5)		
0	D4	0.0115102
1	D4, D3	0.0105300
2	D4, D3, C6	**0.0103634**
3	D4, D3, C6, C2	0.0104309
4	D4, D3, C6, C2, C4	0.0106421
July (D4)		
0	D5	0.0069262
1	D5, F1	0.0062005
2	D5, F1, D6	0.0061786
3	D5, F1, D6, D3	0.0051103
4	D5, F1, D6, D3, C4	0.0050485
5	D5, F1, D6, D3, C4, D2	0.0050180
6	D5, F1, D6, D3, C4, D2, C6	0.0048374
7	D5, F1, D6, D3, C4, D2, C6, F7	0.0047972
8	D5, F1, D6, D3, C4, D2, C6, F7, E1	**0.0046052**
9	D5, F1, D6, D3, C4, D2, C6, F7, E1, H7	0.0046170
10	D5, F1, D6, D3, C4, D2, C6, F7, E1, H7, F4	0.0046221
October (B7)		
0	B1	0.0397534
1	B1, D3	**0.0355036**
2	B1, D3, B2	0.0373681
3	B1, D3, B2, F1	0.0381142

(Continued)

TABLE 2 | (Continued)

NB: *Sensor location in parenthesis implies the location with the least RMSE value - most important (predictor). The bolded values represent optimal sensor locations.*

Optimal Sensors Placement Problem Formulation

Objective 1: Minimizing the RMSE (Sensor Location Ranking)

A single sensor location that gives the maximum gain to the objective function (Equation 6) was selected from all the environment's 56 possible positions. Furthermore, having fixed the previous selection of the best sensor location, the following location was determined from the remaining (56 – 1 = 55) locations that gave the best improvement in the objective – lowest RMSE. This technique was applied iteratively until the last sensor location was determined. That is, $RMSE_{min,1}$, $RMSE_{min,2}$, $RMSE_{min,3}$,. . ., $RMSE_{min,56}$, where 1, 2,. . ., 56 are placeholders for the sensor nodes, A1, A2,. . ., H7 (not necessarily in this order but ranked by the minimum RMSE at each node).

$$RMSE_{min\_N} = \sqrt{\frac{\sum_{i=1}^{n} \left(x_i\left(i,t\right) - \hat{x_i}\left(i,t\right)\right)^2}{n}} \tag{6}$$

Where N was sensor location number, $x_i\left(i,t\right)$ is the actual observation of the climatic variables at location i and time t, $\hat{x_i}\left(i,t\right)$ was the estimated value, and n was the total number of nodes or sensor locations.

N submatrices of the matrix, $\mathbf{A}$ of $m \times n$ representing the data, such that, $\mathbf{A} \in R^{m \times n}$ were derived to represent the observations at each sensor node given that $S_{m,p\rightarrow q} \in R^{m \times p\rightarrow q}$ ∀ m, n, p, q ∈ N {p, q < n}. $p \rightarrow q$ took an element from the start of the column of a particular sensor location to the end of the column (a node was defined at column index p and q-1 with temperature and humidity index, respectively, for a two-in-one sensor). Sub matrix, S of elements a_{ij} where i = 1, 2, . . ., m; j = 1, 2, . . ., q-1 is ordered in a rectangular frame as shown in Equation 7.

$$S = \begin{bmatrix} a_{1p}a_{1q-1}a_{2p}a_{2q-1} \vdots\vdots a_{mp}a_{mq-1} \end{bmatrix}, a_{ij} \in R \tag{7}$$

A supervised learning approach was employed for this study. As such, a response or target variable (climatic variables to be predicted by the input features), y_{t+k} was derived by making a k-step forecast of a column (feature) of the submatrix, S for

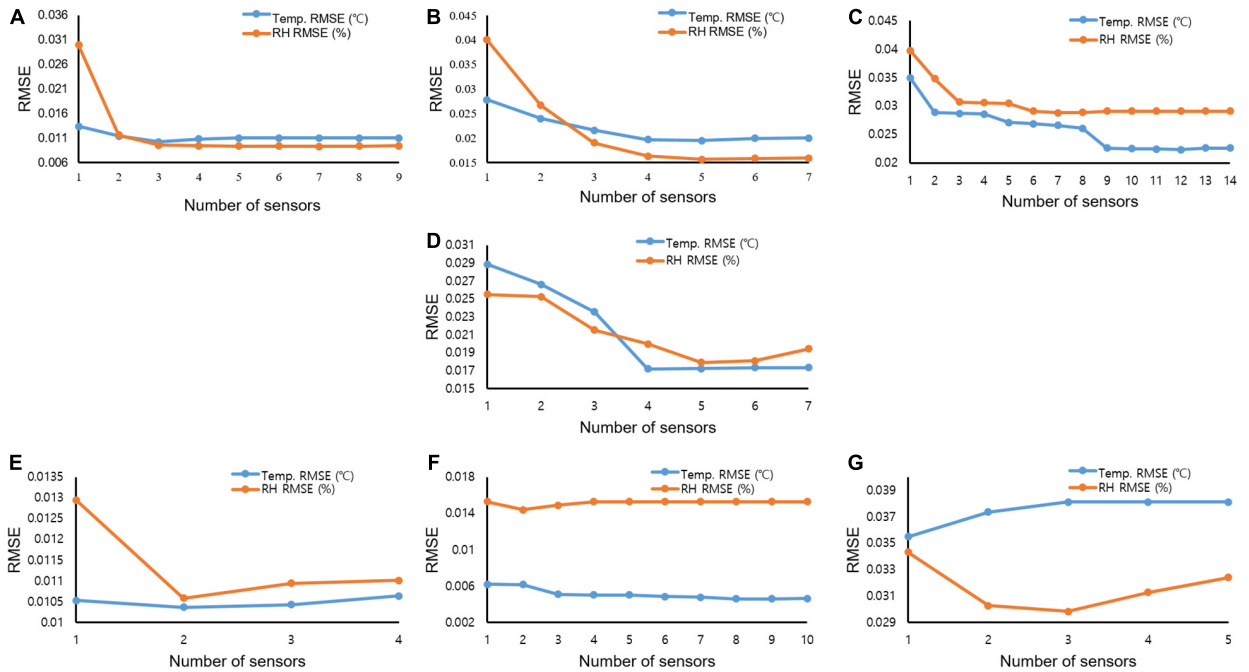


FIGURE 8 | Root mean squared error curves showing the reduction in error at different numbers of sensors using temperature and relative humidity data for **(A)** February; **(B)** March; **(C)** April; **(D)** May; **(E)** June; **(F)** July; and **(G)** October.

$k \in$ N. All observations were made per minute and the response variable at time t was one step ahead of the observation; thus, $k = 1$. A machine learning model simply represented in Equation 8 where d = 1, 2, . . . D were index features, was fitted on the new data matrix, $C \in R^{m \times 3} \supseteq S$ and evaluated by the performance metric, RMSE. This was carried out for all sensor nodes, and the RMSEs were used to rank the order of importance of the sensor nodes – in the order of increasing RMSE values. This implied a larger improvement to the objective function is used to rank the sensors.

$$f : R^D \rightarrow R \tag{8}$$

Objective 2: Minimizing the Optimal Number of Sensors

The second objective of this study was to determine the minimum optimal number of sensors and the sensor location ranking. Having determined the sensor location that gave the most considerable improvement to the objective function, the target variable, y_{t+k} was taken to be a one-step forecast of one of the environmental variables (temperature) readings at this location. Following Li (2016), a gradient boosting model at each point m of M stages, G_m such that $1 \leq m \leq M$ was fitted on the preprocessed data, with the subsequent addition of some estimators, $h_m(x)$(regression trees) to improve the model by compensating for the inadequacy of the existing model $G_m(x)$ (Equation 9).

$$G_{m1}(x) = G_m(x) + h_m(x) \tag{9}$$

$G_{m1}(x)$ is the new model, $G_m(x)$ is the existing model and $h_m(x)$ is the regression tree.

As a supervised learning problem for the training data, $\{(x_1, y_1), ..., (x_n, y_n)\}$, an approximation function, $\hat{G}(x)$ extended a function $G(x)$ to minimize the objective function given as $R(y_{t+k}, G(x))$ by starting with a model containing function $G_0(x)$ and expanding the model as given in Equations 10 and 11.

$$G_0(x) = arg\ \min_{t} \sum_{i=1}^{n} R(y_i, \mu) \tag{10}$$

$$G_m(x) = G_{m-1}(x) + [\sum_{i=1}^{n} R(y_i, G_{m-1}(x_i) + h_m(x_i))] \tag{11}$$

where $i \in N$, $h_m \in H$ is a base learner function.

The model was updated by applying the steepest gradient descent to the minimization problem in Equations 12 and 13.

$$G_m(x) = G_{m-1}(x) - \mu_m \sum_{i=1}^{n} \nabla_{G_{m-1}} R(y_i, G_{m-1}(x_i)) \tag{12}$$

$$\mu_m = \sum_{i=1}^{n} R(y_i, G_{m-1}(x_i) - \mu \nabla_{G_{m-1}} R(y_i, G_{m-1}(x_i))) \tag{13}$$

where the derivatives are taken concerning the functions Gi for $i \in \{1, , m\}$ and $_m$ was the step length.

An N number of sensor locations was considered, where $RMSE_{min,1} < RMSE_{min,2} < RMSE_{min,3} < RMSE_{min,4} < \ldots < RMSE_{min,N}$ for each node from 1, 2, ..., N. A one-step-ahead time forecast at node 1 was taken as the response

TABLE 3 | Performance of sensor network in identifying the optimal number of sensors and placement for measuring greenhouse conditions across different months using relative humidity data.

Index	Sensor location(s)	RMSE (%)
February (F7)		
0	C2	0.0299539
1	C2, H7	0.0115791
2	C2, H7, B7	0.0095410
3	C2, H7, B7, A1	0.0094043
4	C2, H7, B7, A1, D6	0.0093536
5	C2, H7, B7, A1, D6, F1	0.0093313
6	C2, H7, B7, A1, D6, F1, C3	**0.0092886**
7	C2, H7, B7, A1, D6, F1, C3, D3	0.0093494
8	C2, H7, B7, A1, D6, F1, C3, D3, F2	0.0094121
March (G7)		
0	H7	0.0399899
1	H7, D5	0.0266852
2	H7, D5, F7	0.0190371
3	H7, D5, F7, A7	0.0163404
4	H7, D5, F7, A7, B7	**0.0156854**
5	H7, D5, F7, A7, B7, B4	0.0158632
6	H7, D5, F7, A7, B7, B4, C7	0.0159132
April (F7)		
0	D6	0.0397690
1	D6, D7	0.0347859
2	D6, D7, F6	0.0307091
3	D6, D7, F6, F4	0.0305697
4	D6, D7, F6, F4, H7	0.0304953
5	D6, D7, F6, F4, H7, G7	0.0290896
6	D6, D7, F6, F4, H7, G7, F5	**0.0288381**
7	D6, D7, F6, F4, H7, G7, F5, F3	0.0289097
8	D6, D7, F6, F4, H7, G7, F5, F3, E6	0.0291045
May (D4)		
0	A2	0.0255216
1	A2, B2	0.0252681
2	A2, B2, B3	0.0215447
3	A2, B2, B3, F1	0.0199938
4	A2, B2, B3, F1, D3	**0.0179418**
5	A2, B2, B3, F1, D3, H1	0.0181023
6	H7, D5, F7, A7, B7, B4, C7	0.0194512
June (D5)		
0	D4	0.0129441
1	D4, D3	**0.0105844**
2	D4, D3, C6	0.0109450
3	D4, D3, C6, C2	0.0110102
July (D4)		
0	D5	0.0152904
1	D5, F1	**0.0143873**
2	D5, F1, D6	0.0148891
3	D5, F1, D6, D3	0.0152913
October (B7)		
0	B1	0.0343254
1	B1, D3	0.0302872
2	B1, D3, B2	**0.0298292**
3	B1, D3, B2, F1	0.0312717
4	B1, D3, B2, F1, D2	0.0324157

NB: *Sensor location in parenthesis implies the location with the least RMSE value - most important (predictor). The bolded values represent optimal sensor locations.*

or target variable to be used as a predictor for other nodes to determine the performance of placement, while, for the first aspect, the input features were environmental variables (temperature and humidity), and the second aspect, the four psychometric properties (dew point temperature, humid ratio, enthalpy, and specific volume), plus crucially engineered features of the date/time object variable. The overall RMSE continued to decrease, indicating improvements in the sensor stacking performance until a point was reached where there was no further minimization of the objective function. At this point, the number of sensors was considered as being optimal. The pseudo-code below illustrates the algorithm for optimal sensor selection, and **Figure 6** shows the summary of this process. The algorithm used flow conditional statements that iterated the whole process of ranking. The RMSE was the objective function. It was the metric for evaluating the variability due to the disturbances in the greenhouse's climate. The ranking was done by using a time-series forecast methodology. The RMSE compared the predicted values with the actual values.

The RMSE was minimized by tuning the hyper-parameters of the algorithm to obtain the best result. This also ensured that the ranking was not subjected to fluctuations and the optimal selection was accurate no matter how many times the pipeline was automated/re-run.

Optimal sensor selection pseudo-code.

Input: Temperature-humidity dataset, $A \in \mathbb{R}^{m \times n}$, $of (m \times n)$ dimension
Output: Set of optimal sensors
Ranking of. . .
Ranking of sensors
1 Create data-frame for each sensor location
2 Split into X_{train} and X_{val} in a time-series manner
3 Derive target variable by forecasting a feature's observations
4 Split into y_{train} and y_{val}
5 Fit the model on the data
6 Append sensor locations to a list by RMSE ranking (Update the list of sensors iteratively with the corresponding RMSE values in an ascending order to show ranking)
Optimal number of sensors
7 Initialize optimal sensors (p) to 0
8 Assign v to 56 (total number of sensors)
9 for i = 1 to v do
10 Fit model on training set and score
11 if hyper-parameter is not optimal do
12 try other combinations of hyper-parameters
13 else do
14 append RMSE values [Equation 6]
15 increment p by 1
16 return p

EXPERIMENTAL RESULTS AND DISCUSSION

In (Lee et al., 2019), a statistical approach was adopted for optimal sensor selection. Our study advanced the optimal sensor selection by developing a machine learning model using generated time-series big data and transformed psychrometric variables. In

TABLE 4 | Seasonal variation in optimal sensor placement.

February (Winter)					March (Spring)					April (Spring)					May (Spring)					June (Summer)					July (Summer)					October (Autumn)				
U	Td	w	h	v	U	Td	w	h	v	U	Td	w	h	v	U	Td	w	h	v	U	Td	w	h	v	U	Td	w	h	v	U	Td	w	h	v
F7	F7	F7	F7	F7	G7	G7	G7	G7	G7	F7	F7	F7	F7	F7	D4	D4	D4	D4	D4	D5	D5	D5	D5	D5	D4	D4	D4	D4	D4	B7	B7	B7	B7	B7
C2	C2	C2	C2	C2	H7	H7	H7	H7	H7	D6	D6	D6	D6	D6	A2	A2	A2	A2	A2	D4	D4	D4	D4	D4	D5	D5	D5	D5	D5	B1	B1	B1	B1	B1
H7	H7	H7	H7	H7	D5	D5	D5	D5	D5	D7	D7	D7	D7	D7	B2	B2	B2	B2	B2	D3	D3	D3	D3	D3	F1	F1	F1	F1	F1	D3	D3	D3	D3	D3
		B7	B7	B7	F7	F7	F7	F7		F6	F6	F6	F6	F6	B3	B3	B3	B3	B3	C6	C6	C6	C6	C6	D6	D6	D6	D6	D6					
			A1	A1	A7	A7	A7	A7		F4		F4	F4	F4	F1	F1	F1	F1	F1	C2					D3	D3	D3	D3	D3					
			D6	D6	B7	B7	B7	B7		H7		H7	H7	H7	D3		D3	D3		C4					C4	C4	C4	C4	C4					
					B4	B4	B4			G7		G7	G7	G7	H1		H1	H1							D2	D2	D2	D2	D2					
					C7	C7	C7			F5		F5			A4		A4	A4							C6	C6	C6	C6	C6					
					E7	E7	E7			F3		F3			C4		C4	C4							F7	F7	F7	F7	F7					
						D7	D7			E6		E6			A3		A3	A3								E1	E1	E1	E1					
						D4	D4			E7		E7			D2											H7								
										H6		H6			C2											F4								
										D2		D2			E1											F5								
												D5			C1																			
															A6																			
															C3																			
															B1																			
(3)	(3)	(6)	(6)	(4)	(9)	(11)	(11)	(6)	(3)	(13)	(4)	(14)	(7)	(7)	(17)	(5)	(10)	(10)	(5)	(6)	(4)	(4)	(4)	(4)	(9)	(13)	(10)	(10)	(10)	(3)	(3)	(3)	(3)	(3)

Keys: U – untransformed (raw temperature and humidity) data. Td – dew point temperature. w – humidity ratio. h – enthalpy. v - specific volume.

the results obtained in Lee et al. (2019), sensor locations with the highest entropy were selected as optimal because of high disturbance from the wind. We implemented an algorithm on time-series big data and transformed psychrometric variables that choose sensors that can best monitor the state of the greenhouse optimally using hyper-parameter tuning.

Variability Analysis: Coefficient of Variation of Greenhouse Temperature and Humidity Data

The coefficient of variation (*CV*) of the climate data was calculated for all the studied months. The *CV* was used to determine the extent of variability of the greenhouse by computing the ratio of the standard deviation to the mean of the temperature or relative humidity values.

During the summer period (June and July) in **Table 1**, it was observed that the temperature variation was the least, indicating the data points have the minimum difference from the mean compared to other periods. July showed the least CV for the relative humidity data but differed slightly from June and showed a slightly higher value than February. Similarly, the greenhouse had the least variability in the summer months for relative humidity. Considering plants were not grown during this period in the greenhouse could be a reason for the low variation in greenhouse climate properties. During aerobic respiration, plants use oxygen and emit carbon dioxide (Kader and Saltveit, 2002), which affects the properties of the greenhouse. Generally, this was observed in other months in which plants were grown.

February and March (the end of winter and the beginning of spring) are the two months with the highest temperature *CV* (40.43% and 42.30%, respectively), similarly with a very high humidity *CV* (32.53% and 42.09%, respectively). This was probably caused by the changing season, with a sharp change in weather conditions. February and March recorded a low of −9°C and −2°C, respectively. Both months had a high of 24°C. An increment of 9.15% and 30.94% in the temperature and humidity standard deviation, respectively, were observed in March.

The Effect of the Plants on the Microclimate Distribution

The dynamic time warping algorithm was implemented to measure the similarity between March and June sensor readings. **Figure 7A** shows the plot of the March and June sensor reading per hour for the temperature data, and **Figure 7B** shows the alignment match plot of the series. The optimal match between the two series, as shown in **Figure 7B**, cannot be understood visually since the dataset is quite large.

The first index from the *i* sequence matches with at least 250 indices of the *j* sequence. This implies that the first hour of June matches with the first 11 days of March with a minimal cost path as indicated by the vertical line. This implied that there were no statistically significant changes in the climatic condition of the greenhouse. However, before the end of the first day in June, a slow change in the graph indicated a shift of alignment between the two months. A significant match occurred at the 170th index of the *j* sequence (7th day of June) with the 352^{nd} index of the *i* sequence (toward the evening of the 14th day of March).

This similarity was stable for about 48 h (2 days). The overall climatic condition of the greenhouse, by the temperature, in the last seven days in June matched closely with the state of the greenhouse within the previous two days of March, with minor variations.

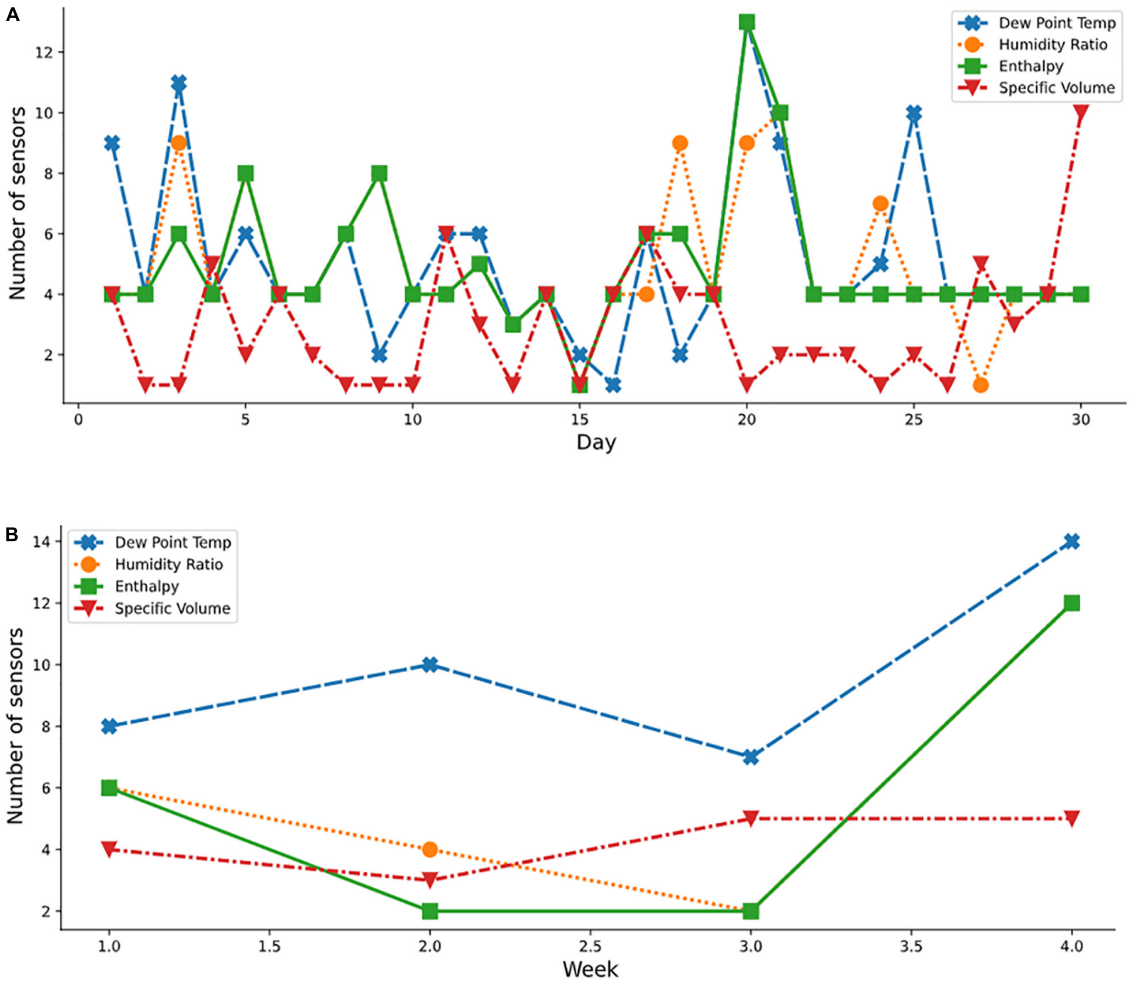


FIGURE 9 | Optimal sensor selection for spring month using the psychometric dataset; **(A)** April daily; and **(B)** weekly.

Almost all points of the *i* and *j* indices had unique matches for the relative humidity data. A lesser number of matches of the indices was observed. This high linearity implied that the absence of crops in June did not have much effect on the relative humidity of the greenhouse compared to the temperature. This also justified a lesser percentage decrease from March to June in the relative humidity *CV* than the *CV* of temperature as shown in **Table 1**.

Temperature-Relative Humidity Data

The optimal hyper-parameters were *n_estimators* = 1000 and *learning_rate* = 0.01, while *max_depth* ranged from 2 to 7, as selected by the iterative algorithm. These hyper-parameters were used to tune the algorithms to learn the data with the maximum performance. Seven months (February, March, April, May, June, July, and October) were selected as representative months to cover the four seasons (winter, spring, summer, and autumn) and used in the simulations.

In **Table 2**, index numbers 3, 5, 12, 4, 3, 8, and 1 with the least RMSE values of 0.0102077, 0.0194982, 0.0223317, 0.0171915, 0.0103634, 0.0046052, and 0.0355036 were recorded as optimal sensors numbers and locations for February, March, April, May, June, July, and October, respectively, for temperature data. At some months, the RMSE values start increasing, indicating that the addition of more sensors would instead reduce the quality of the data. These presented the sensors that measured the air-moisture condition in the greenhouse most accurately in the different months. The sensors acted as features or variables used for training the machine learning model. Through ranking, the number of sensors required was determined with the RMSE indicating the model's performance in predicting the best sensor location. The more relevant the feature(s), the lower the RMSE. The row (bolded) beyond which the RMSE no longer decreased was taken as the optimal. **Table 2** shows index number that a high variation in the optimal number of sensors occurred at different months with a total number of 4, 6, 13, 5, 3, 9, and 3 sensors were optimal for measuring the greenhouse's internal environment in February, March, April, May, June, July, and October, respectively.

Furthermore, investigation of the Pareto front, a set of nondominated solutions chosen as optimal when no objective can be improved without sacrificing at least one other objective,

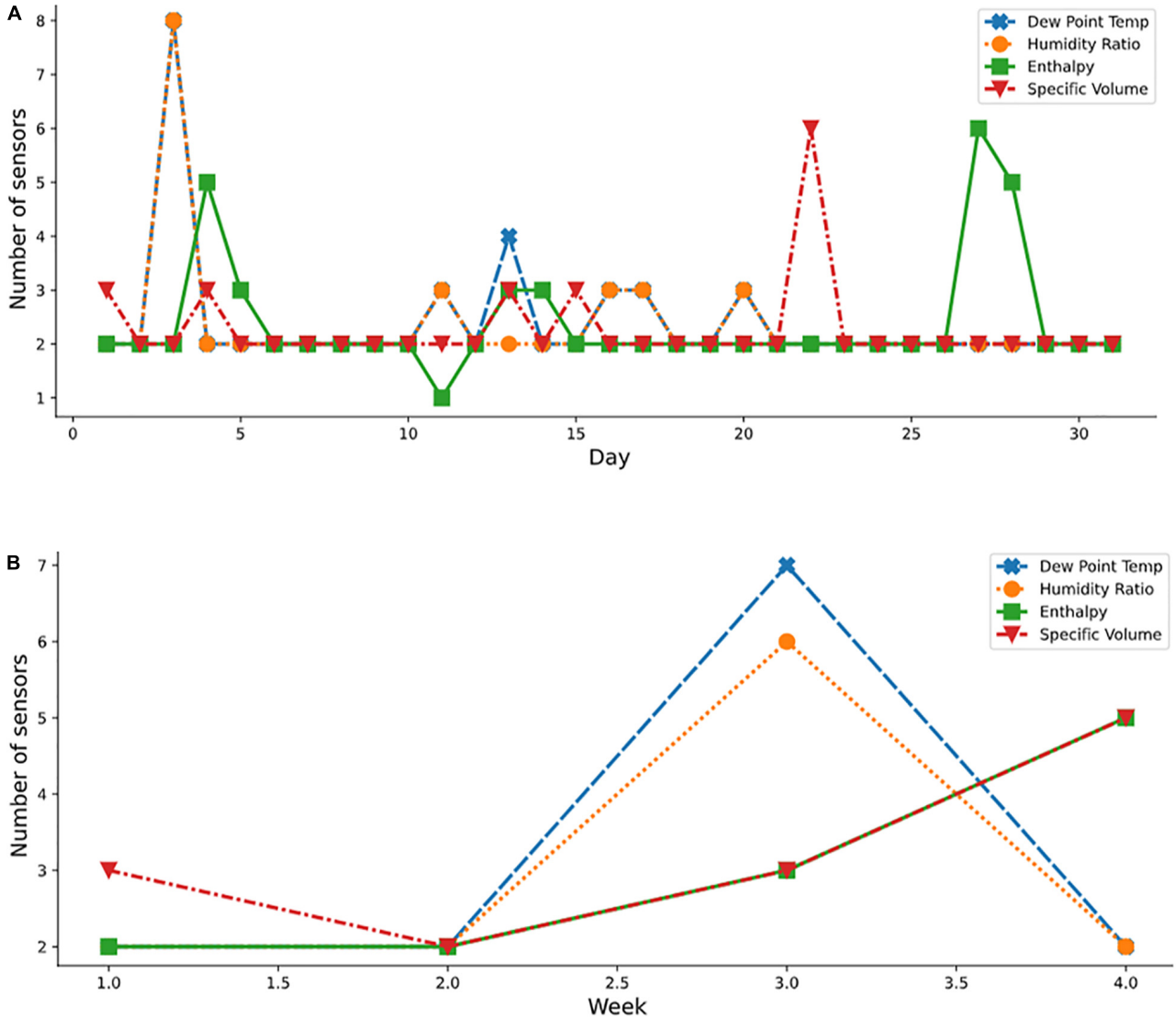


FIGURE 10 | Optimal sensor selection for summer month using the psychometric dataset; **(A)** July daily; and **(B)** weekly.

helped enhance decision-making. The two conflicting objectives showed a reduction in RMSE values at all the investigated months with increasing selected sensors. However, to reduce the RMSE and number of conflicting sensors, the Pareto front displayed the knee points where a less significant RMSE occurred. In February, (**Figure 8A**), a drastic reduction (about 66%) in the RMSE value between one and two sensors with a slighter decrease between two and five sensors using the temperature data. This indicated that, for February, two sensors would give good readings to understand the condition of the air-vapor mixture in the greenhouse at a less computational cost than three (about 0.014%), four, and five sensors (about 13%). This trend was seen for the other simulated months, with March (**Figure 8B**) having three knee points at two and five sensors with about 26% and 46%, respectively. In April (**Figure 8C**), two distinct knee points were recorded at three sensors (about 28%) and ten sensors (about 42%). A similar trend was seen in May (**Figure 8D**), June (**Figure 8E**), and July (**Figure 8F**). However, in October, a drastic reduction was seen at two sensors (about 30%), followed by a sharp rise indicating that more sensors introduced more errors instead (**Figure 8G**).

Index numbers 6, 4, 6, 4, 1, 1, and 2 with the least RMSE values of 0.0092886, 0.0156854, 0.0288381, 0.0179418, 0.0105844, 0.0143873, and 0.0298292 were recorded as optimal sensors numbers and locations for February, March, April, May, June, July, and October, respectively, as the sensors that measured the air-moisture condition in the greenhouse most accurately in the other months using the relative humidity data. The results for the sensors to measure the air-vapor mixture in the greenhouse differed from the temperature and humidity data. This led us to investigate the stability of the transformed data

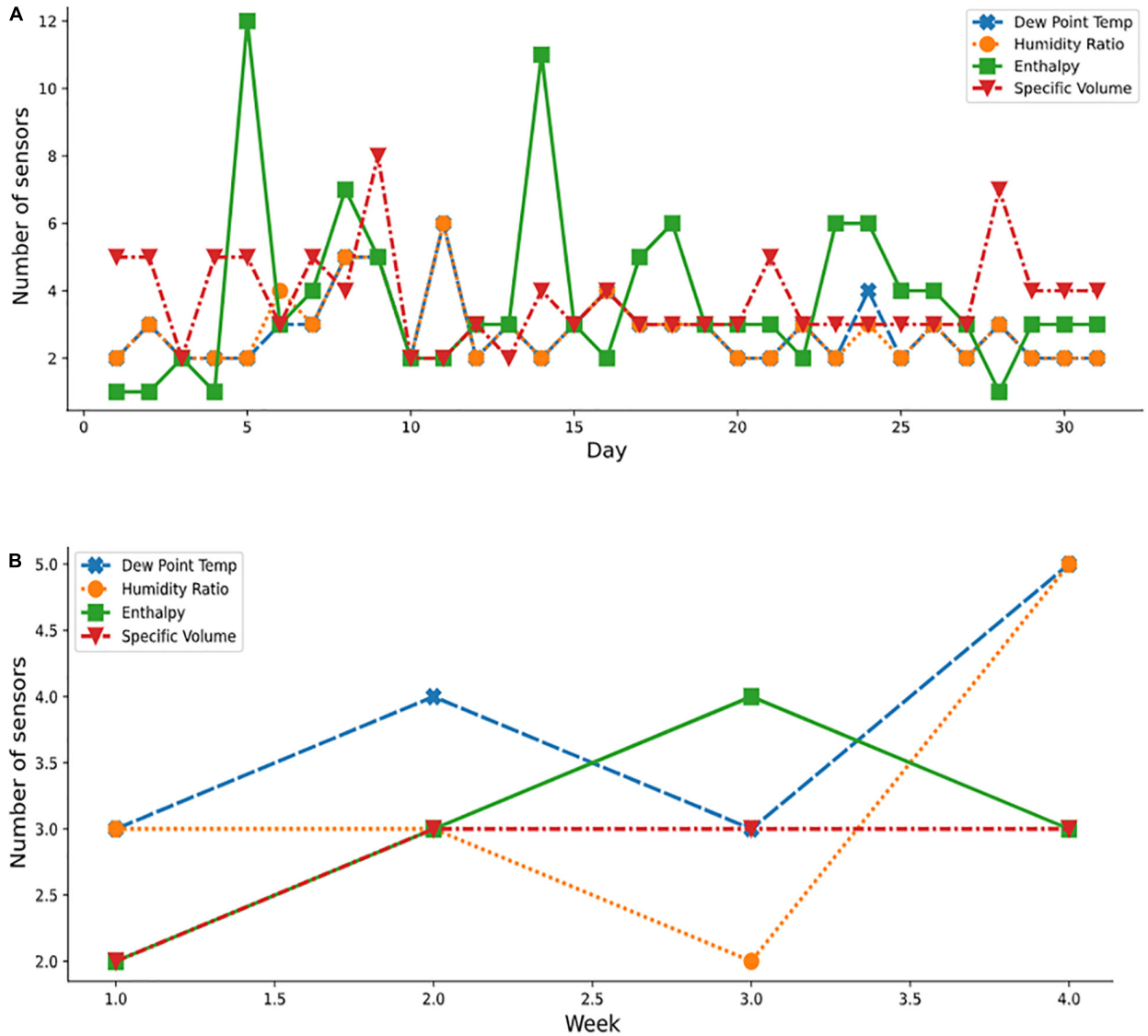


FIGURE 11 | Optimal sensor selection for Autumn month using the psychometric dataset; **(A)** October daily; and **(B)** weekly.

would best describe the air-vapor mixture condition in the greenhouse. In the case of relative humidity, a similar trend of high variation in the optimal number of sensors occurred at different months, with a total number of 7, 5, 7, 5, 2, 2, and 3 sensors being found optimal for measuring the internal greenhouse environment in February, March, April, May, June, July, and October, respectively (**Table 3**).

A similar trend with the temperature data was seen in the Pareto fronts for the optimal number of sensors to accurately measure the air-vapor mixture in the greenhouse using the relative humidity data. A flat Pareto front was seen in the reduction of RMSE in February (**Figure 8A**) from 2 to 8 sensors. All other months (**Figures 8B–G**) showed that the Pareto front improved decision-making, as there were sensors that though reduced the RMSE, did not significantly cause the front to change.

The temperature and relative humidity data had the same predictor, implying the location with the least RMSE value (**Tables 2**, **3**). This indicated that the rankings were not different (shows the sensors with the least interference for the month). However, the optimal numbers of sensors using the temperature and relative humidity data varied across the months.

Transformed Data: Psychometric Variables

The optimal sensor locations for February, March, April, May, June, July, and October are given in **Table 4** for the four psychrometric properties considered in this study. The monthly data was split into daily and weekly data to get a clearer view of optimal sensors placement for each month and investigate the effect of the sharp changes in weather conditions. Analyses of the

sensor numbers results show that the transformed psychrometric variables had fewer optimal locations than the untransformed (temperature and relative humidity) dataset, with a difference of up to about 70% in May.

Figures 9A,B show the daily and weekly distributions of the variables by sensors for a spring month (April), respectively. Using the transformed psychometric properties (dew point temperature, humid ratio, enthalpy, and specific volume), inconsistencies in selections of the ideal number of sensors required for some days in April were observed. For instance, on day 1 (April), nine optimal sensors were required when the dew point temperature property was considered, while four optimal sensors were required for other properties. The usage of the derived psychometric properties resulted in the selection of a reduced number of optimal sensors indicating a more adaptive nature of the algorithm to these derived variables compared to the raw temperature and relative humidity variables. The derived psychometric properties also showed better understanding of the air-vapor mixture since two combined properties were considered instead of the untransformed dataset using a single property. This improved efficiency would benefit the grower by reducing acquisition and operating costs as well as decreasing amounts of dat ato be handled. Furthermore, a cross-cutting beneficial effect would result from energy savings from proper monitoring and increased productivity.

Over the study period, specific volume (v) required the least number of sensors for measurement. However, it showed the most inconsistent result (having values not in a close range with the result from other derived properties), likely due to the very low magnitude of the values producing slightly more stochastic predictions.

Additionally, it was noted that the order of sensor selection did not change over the study period. For example, as reported for April in **Table 4**, in the April column, the ranking of the 13 sensor locations according to decreasing order of importance was F7, D6, D7, F6, F4, H7, G7, F5, F3, E6, E7, H6, and D2. If four sensors were required for measuring the enthalpy variable, then the first four sensor locations (F7, D6, D7, and F6) were to be considered. If one sensor only was selected, then F7 was the optimal sensor. **Figures 10**, **11** show the periodic variation of optimal sensor selection for summer (July) and autumn (October), respectively.

Periodic Variation in the Optimal Sensor Selection

Several plots (**Figures 9–11**) show the variation in the optimal sensor selection for the greenhouse over time (that is, daily and weekly). External disturbances such as temperature, wind, and humidity influenced the data. Modeling the phenomenon of natural ventilation proved to be complex, especially because it was significantly affected by the external climate, and its design more complicated than fan ventilation. A fan ventilation system was adopted for verification (**Figure 3C**). Yet, significant variations were still observed based on the analysis of the coefficient of variation of the indoor climate data (**Table 1**). Also thought to be influencing the microclimate within the protected cultivation systems were factors such as the heating system and the respiration of the plants which could have led to variations in the relative humidity. This necessitated a systematic approach for determining the optimal number and locations of the sensors. For example, on days 15 and 20 of April (**Figure 9A**), temperature measurements and standard deviations of 0.3258°C and 0.2130°C, respectively were reported. Statistically, in terms of measuring dispersion, the magnitude of the standard deviation varied across daily, weekly, and monthly periods. These temporal variations in the results indicated that optimal sensor placement was affected by periodic variations of different levels of magnitude. However, the sensors selected at the same level across the measured conditions and transformed psychrometric properties were the same, pointing to the robustness of our method to accurately measure the total air-vapor mixture in the protected cultivation system.

Additionally, the indoor heating system could contribute to the differences in the optimal locations selected to measure environmental conditions as some piping systems heat sections of the greenhouse nonuniformly.

CONCLUSION

A supervised machine learning model was developed to identify the optimal number and locations of sensors to monitor climatic conditions in a protected cultivation system using a multi-objective approach. The Gradient Boosting Algorithm was fitted to the measured conditions and derived psychrometric variables. The derived psychrometric properties resulted in fewer optimal sensors than the raw temperature and relative humidity data. This study found that the optimal locations of sensors were both at the sides and center of the protected cultivation system depending on the time of year. Variability analyses indicated that no location was consistently optimal. The changes in the optimal sensor location with seasons were this study's limitation. A future study would aim to develop a dynamic approach to selecting optimal sensors' locations. This could include using the ensemble technique by creating multiple models and considering a mobile environmental measurement system. Finally, the solutions in the Pareto front improved decision-making as some points had close relationships. This would have cross-cutting effects on energy management and plant productivity.

DATA AVAILABILITY STATEMENT

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

AUTHOR CONTRIBUTIONS

DU: conceptualization, methodology, investigation, formal analysis, data curation, supervision, visualization, and writing – original draft. OI: methodology, investigation, software,

data curation, visualization, and writing – original draft. RM: methodology, investigation, software, data curation, visualization, and writing – review and editing. SA-H: methodology, validation, data curation, visualization, supervision, and writing – review and editing. YH: validation, resources, writing – review and editing, and supervision. MA: methodology, investigation, and supervision. TP: methodology, investigation, software, data curation, visualization, validation, resources, writing – review and editing, supervision, and funding acquisition. All authors contributed to the article and approved the submitted version.

FUNDING

This work was supported by Korea Institute of Planning and Evaluation for Technology in Food, Agriculture, and Forestry (IPET) through Agriculture, Food and Rural Affairs Convergence Technologies Program for Educating Creative Global Leader, funded by Ministry of Agriculture, Food and Rural Affairs (MAFRA) (320001-4) and (716001-7), South Korea and the Basic Science Research Program through the National Research Foundation of Korea (NRF) through the Ministry of Education under Grant 2020R1I1A1A01073794.

REFERENCES

Arnesano, M., Revel, G., and Seri, F. (2016). A tool for the optimal sensor placement to optimize temperature monitoring in large sports spaces. *Autom. Constr.* 68, 223–234. doi: 10.1016/j.autcon.2016.05.012

Ayalew, D., Tesfaye, K., Mamo, G., Yitaferu, B., and Bayu, W. (2012). Variability of rainfall and its current trend in Amhara region, Ethiopia. *Afr. J. Agric. Res.* 7, 1475–1486. doi: 10.5897/AJAR11.698

Aydin, B. E., Hagedooren, H., Rutten, M. M., Delsman, J., Oude Essink, G. H., van de Giesen, N., et al. (2019). A greedy algorithm for optimal sensor placement to estimate salinity in polder networks. *Water* 11:1101. doi: 10.3390/w11051101

Bhujel, A., Basak, J. K., Khan, F., Arulmozhi, E., Jaihuni, M., Sihalath, T., et al. (2020). Sensor systems for greenhouse microclimate monitoring and control: a review. *J. Biosyst. Eng.* 45, 341–361. doi: 10.1007/s42853-020-00075-6

Boerema, S. T., Van Velsen, L., Schaake, L., Tönis, T. M., and Hermens, H. J. (2014). Optimal sensor placement for measuring physical activity with a 3D accelerometer. *Sensors* 14, 3188–3206. doi: 10.3390/s140203188

Chang, D.-E., Ha, K.-R., Jun, H.-D., and Kang, K.-H. (2012). Determination of optimal pressure monitoring locations of water distribution systems using entropy theory and genetic algorithm. *J. Korean Soc. Water Wastewater* 26, 1–12. doi: 10.11001/jksww.2012.26.1.001

Chang, M., and Pakzad, S. N. (2014). Optimal sensor placement for modal identification of bridge systems considering number of sensing nodes. *J. Bridge Eng.* 19:04014019. doi: 10.1061/(ASCE)BE.1943-5592.0000594

Cossu, M., Murgia, L., Ledda, L., Deligios, P. A., Sirigu, A., Chessa, F., et al. (2014). Solar radiation distribution inside a greenhouse with south-oriented photovoltaic roofs and effects on crop productivity. *Appl. Energy* 133, 89–100. doi: 10.1016/j.apenergy.2014.07.070

Czubinski, F. F., Mantelli, M. B., and Passos, J. C. (2013). Condensation on downward-facing surfaces subjected to upstream flow of air–vapor mixture. *Exp. Therm. Fluid Sci.* 47, 90–97. doi: 10.1016/j.expthermflusci.2013.01.004

DeFacio, P., Pickerel, L., and Rhyne, S. M. (2002). *Greenhouse Operation and Management: Instructional Materials Laboratory*. Columbia, MO: University of Missouri, 10.

Du, W., Xing, Z., Li, M., He, B., Chua, L. H. C., and Miao, H. (2014). "Optimal sensor placement and measurement of wind for water quality studies in urban reservoirs. IPSN-14," in *Proceedings of the 13th International Symposium on Information Processing in Sensor Networks*, (Piscataway, NJ: IEEE). doi: 10.1109/IPSN.2014.6846750

Faris, D. M., and Mahmood, M. B. (2014). Data acquisition of greenhouse using Arduino. *J. Babylon Univ.* 22, 1908–1906.

Feng, L., Li, H., and Zhi, Y. (2013). "Greenhouse CFD simulation for searching the sensors optimal placements," in *Proceedings of the 2013 Second International Conference on Agro-Geoinformatics (Agro-Geoinformatics)*, (Piscataway, NJ: IEEE). doi: 10.1109/Argo-Geoinformatics.2013.6621972

Fontanini, A. D., Vaidya, U., and Ganapathysubramanian, B. (2016). A methodology for optimal placement of sensors in enclosed environments: a dynamical systems approach. *Build. Environ.* 100, 145–161. doi: 10.1016/j.buildenv.2016.02.003

Furlanello, C., Merler, S., and Jurman, G. (2006). Combining feature selection and DTW for time-varying functional genomics. *IEEE Trans. Signal Process.* 54, 2436–2443. doi: 10.1109/TSP.2006.873715

Gadekallu, T. R., Rajput, D. S., Reddy, M., Lakshmanna, K., Bhattacharya, S., Singh, S., et al. (2021). A novel PCA–whale optimization-based deep neural network model for classification of tomato plant diseases using GPU. *J. Real Time Image Process.* 18, 1383–1396. doi: 10.1007/s11554-020-00987-8

Graamans, L., Baeza, E., Van Den Dobbelsteen, A., Tsafaras, I., and Stanghellini, C. (2018). Plant factories versus greenhouses: comparison of resource use efficiency. *Agric. Syst.* 160, 31–43. doi: 10.1016/j.agsy.2017.11.003

Guzmán, C. H., Carrera, J. L., Durán, H. A., Berumen, J., Ortiz, A. A., Guirette, O. A., et al. (2019). Implementation of virtual sensors for monitoring temperature in greenhouses using CFD and control. *Sensors* 19:60. doi: 10.3390/s19010060

Handbook, A. (2001). *Fundamentals, 2001*. Atlanta: ASHRAE.

Hemming, S., Zwart, F. D., Elings, A., Petropoulou, A., and Righini, I. (2020). Cherry tomato production in intelligent greenhouses—sensors and AI for control of climate, irrigation, crop yield, and quality. *Sensors* 20:6430. doi: 10.3390/s20226430

Hu, J., and Patel, M. (2014). "Optimized selection and placement of sensors using building information models (BIM)," in *Proceedings of the IES Annual Conference*, Pittsburgh, PA.

Huang, G., Zhou, P., and Zhang, L. (2014). "Optimal location of wireless temperature sensor nodes in large-scale rooms," in *Proceedings of the 13th International Conference on Indoor Air Quality and Climate, Indoor Air*, Hong Kong.

Jones, M. B. (1985). "Chapter 3 – plant microclimate," in *Techniques in Bioproductivity and Photosynthesis*, 2nd Edn, eds J. Coombs, D. O. Hall, S. P. Long, and J. M. O. Scurlock (Oxford: Pergamon Press), 26–40. doi: 10.1016/B978-0-08-031999-5.50013-3

Kader, A. A., and Saltveit, M. E. (2002). *Respiration And Gas Exchange: Postharvest Physiology And Pathology Of Vegetables*. Boca Raton, FL: CRC Press, 31–56. doi: 10.1201/9780203910092.ch2

Kassie, B. T. (2014). *Climate variability And Change In Ethiopia: Exploring Impacts And Adaptation Options For Cereal Production*. Wageningen: Wageningen University and Research.

Lee, S.-Y., Lee, I.-B., Yeo, U.-H., Kim, R.-W., and Kim, J.-G. (2019). Optimal sensor placement for monitoring and controlling greenhouse internal environments. *Biosyst. Eng.* 188, 190–206. doi: 10.1016/j.biosystemseng.2019.10.005

Li, C. (2016). *A Gentle Introduction To Gradient Boosting*. Available online at: https://www.ccs.neu.edu/home/vip/teach/MLcourse/4_boosting/slides/gradient_boosting (accessed December 13, 2020).

Li, S., Cheng, X., Chen, Y., and Zhang, H. (2013). The optimal placement of sensors in square target regions with varying boundary length. *Proc. Eng.* 62, 899–906. doi: 10.1016/j.proeng.2013.08.141

Li, X.-H., Cheng, X., Yan, K., and Gong, P. (2010). A monitoring system for vegetable greenhouses based on a wireless sensor network. *Sensors* 10, 8963–8980. doi: 10.3390/s101008963

Löhner, R., and Camelli, F. (2005). Optimal placement of sensors for contaminant detection based on detailed 3D CFD simulations. *Eng. Comput.* 22, 260–273. doi: 10.1108/02644400510588076

Nelson, P. V. (1991). *Greenhouse Operation And Management*. Hoboken, NJ: Prentice Hall.

Pamungkas, A. P., Hatou, K., and Morimoto, T. (2014). Evapotranspiration model analysis of crop water use in plant factory system. *Environ. Control Biol.* 52, 183–188. doi: 10.2525/ecb.52.183

Park, D.-H., and Park, J.-W. (2011). Wireless sensor network-based greenhouse environment monitoring and automatic control system for dew condensation prevention. *Sensors* 11, 3640–3651. doi: 10.3390/s110403640

Park, S.-H., Park, T., Park, H. D., Jung, D.-H., and Kim, J. Y. (2019). Development of wireless sensor node and controller complying with communication Interface standard for smart farming. *J. Biosyst. Eng.* 44, 41–45. doi: 10.1007/s42853-019-00001-5

Ponce, P., Molina, A., Cepeda, P., Lugo, E., and MacCleery, B. (2014). *Greenhouse Design And Control*. Boca Raton, FL: CRC Press. doi: 10.1201/b17391

Prieto, I., Armas, C., and Pugnaire, F. I. (2012). Water release through plant roots: new insights into its consequences at the plant and ecosystem level. *New Phytol.* 193, 830–841. doi: 10.1111/j.1469-8137.2011.04039.x

Rao, A. R. M., Lakshmi, K., and Krishnakumar, S. (2014). A generalized optimal sensor placement technique for structural health monitoring and system identification. *Proc. Eng.* 86, 529–538. doi: 10.1016/j.proeng.2014.11.077

Seabrook, T. (2016). *Optimal Placement Strategies of Minimum Effective Sensors for Application in Smart Buildings*. Available online at: https://www.semanticscholar.org/paper/Optimal-Placement-Strategies-of-Minimum-Effective-Seabrook/ [Accessed December 10, 2018].

Stanghellini, C. (2013). "Horticultural production in greenhouses: efficient use of water," in *Proceedings of the International Symposium on Growing Media and Soilless Cultivation*, Leuven, 1034. doi: 10.17660/ActaHortic.2014.1034.1

Syed, A. M., and Hachem, C. (2019a). Review of design trends in lighting, environmental controls, carbon dioxide supplementation, passive design, and renewable energy systems for agricultural greenhouses. *J. Biosyst. Eng.* 44, 28–36. doi: 10.1007/s42853-019-00006-0

Syed, A. M., and Hachem, C. (2019b). Review of construction; geometry; heating, ventilation, and air-conditioning; and indoor climate requirements of agricultural greenhouses. *J. Biosyst. Eng.* 44, 18–27. doi: 10.1007/s42853-019-00005-1

Tong, K., Bakhary, N., Kueh, A., and Yassin, A. (2014). Optimal sensor placement for mode shapes using improved simulated annealing. *Smart Struct. Syst.* 13, 389–406. doi: 10.12989/sss.2014.13.3.389

Uyeh, D. D., Pamulapati, T., Mallipeddi, R., Park, T., Asem-Hiablie, S., Woo, S., et al. (2019). Precision animal feed formulation: an evolutionary multi-objective approach. *Anim. Feed Sci. Technol.* 256:114211. doi: 10.1016/j.anifeedsci.2019.114211

Uyeh, D. D., Pamulapati, T., Mallipeddi, R., Park, T., Woo, S., Lee, S., et al. (2021). An evolutionary approach to robot scheduling in protected cultivation systems for uninterrupted and maximization of working time. *Comput. Electron. Agric.* 187:106231. doi: 10.1016/j.compag.2021.106231

Vox, G., Teitel, M., Pardossi, A., Minuto, A., Tinivella, F., and Schettini, E. (2010). *Sustainable Greenhouse Systems: Sustainable Agriculture: Technology, Planning And Management*. New York, NY: Nova Science Publishers Inc, 1–79.

Wang, X., Ma, J., and Wang, S. (2009). Parallel energy-efficient coverage optimization with maximum entropy clustering in wireless sensor networks. *J. Parallel Distrib. Comput.* 69, 838–847. doi: 10.1016/j.jpdc.2009.04.012

Wang, X.-H., Xu, L.-H., and Wei, R.-H. (2014). "A new fusion structure model on greenhouse environment data and a new fusion algorithm of sunlight," in *Proceedings of the 2014 International Conference on Wireless Communication and Sensor Network*, (Piscataway, NJ: IEEE). doi: 10.1109/WCSN.2014.91

Worden, K., and Burrows, A. (2001). Optimal sensor placement for fault detection. *Eng. Struct.* 23, 885–901. doi: 10.1016/S0141-0296(00)00118-8

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

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

frontiers | Frontiers in Plant Science

ORIGINAL RESEARCH
published: 14 July 2022
doi: 10.3389/fpls.2022.898183

Distributed Averaging Problems of Agriculture Picking Multi-Robot Systems *via* Sampled Control

Fengying Ma, Hui Yao, Mingjun Du, Peng Ji and Xiaoan Si*

School of Information and Automation Engineering, Qilu University of Technology (Shandong Academy of Sciences), Jinan, China

Distributed control of agriculture picking multi-robot systems has been widely used in the field of smart agriculture, this paper aims to explore the distributed averaging problems of agriculture picking multi-robot systems under directed communication topologies by taking advantage of the sampled data. With the algebraic graph theory concepts and the matrix theory, a distributed protocol is proposed based on the nearest sampled neighbor information. It is shown that under the proposed protocol, the states of all agents can be guaranteed to reach average consensus whose value is the averaging of the initial states of all agents. Besides, when considering time-delay, the other distributed protocol is constructed, in which a time margin of the time-delay can be determined simultaneously. The necessary and sufficient consensus results can be developed even though the time delay exists. Simulation results are given to demonstrate the effectiveness of our developed consensus results.

Keywords: average consensus, directed communication topologies, distributed protocol, smart agriculture, sampled data

OPEN ACCESS

Edited by:
Lei Shu,
Nanjing Agricultural University, China

Reviewed by:
Saeed Hamood Alsamhi,
Ibb University, Yemen
Mingyang Lu,
The University of Manchester, United Kingdom

****Correspondence:***
Mingjun Du
dumingjun0421@163.com

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

Received: *17 March 2022*
Accepted: *11 July 2022*
Published: *14 July 2022*

Citation:
Ma F, Yao H, Du M, Ji P and Si X (2022) Distributed Averaging Problems of Agriculture Picking Multi-Robot Systems via Sampled Control. Front. Plant Sci. 13:898183. doi: 10.3389/fpls.2022.898183

1. INTRODUCTION

Recently, the issue of smart agriculture has attracted wide attention. Driven by the digital revolution, agriculture has entered a new era of digital and intelligent development (Yang et al., 2013; Alsamhi et al., 2019a; Horng et al., 2019; Fuentes et al., 2021; Li and Chao, 2022; Teng et al., 2022). Smart agriculture is a modern agricultural production mode with information, theoretical knowledge, and hardware equipment as the core elements, and it is an important direction of the development of modern agriculture (Chen and Yang, 2019). Realizing precision agriculture is a goal of smart agriculture (Luo et al., 2016), the present stage of agriculture at a relatively low level of agricultural mechanization, especially in the area of vegetable picking. The traditional way of vegetables is picked manually, which requires a lot of labor during the picking season. Although some plantations have begun to mechanize agricultural picking, the level of automation is low, and vegetable picking is usually carried out by a single mechanical equipment (Brondino et al., 2021), which is inefficient and costly. Therefore, it is urgent to improve the efficiency and mechanization of vegetable picking agriculture.

The emergence of the multi-agent system provides a new trend for the development of smart agriculture. Multi-agent systems are composed of some agents and interactions among agents to solve problems that are impossible for a single agent, which can be applied in many aspects, such as multi-robots (Yu et al., 2020), unmanned aerial vehicles (Lian and Deshmukh, 2006; Alsamhi et al., 2021; Liang et al., 2021), unmanned

ground vehicles (Ma et al., 2006), etc. To realize the cooperation between the agents, communication is the fundamental problem, many scholars have carried out a lot of research on communication. In Alsamhi et al. (2019b), the development status of artificial intelligence in the field of the communication among robots is reviewed, and new ideas for its future research directions are provided. In order to better realize the cooperation between robots and ensure effective communication between robots, Alsamhi et al. (2020) fully analyze and discusses robots in different spatial positions through the fusion of machine learning and communication. A sufficient overview is provided for the use of various machine learning techniques in the communication among robots, and it is shown that machine learning plays an important role in improving the communication among robots. In practical application, information communication will be delayed, which may result in the instability of the control system (Seuret et al., 2018). In Zhang et al. (2022), a more general communication mode between delay subsystems and studies of the effect of time delay on the performance of interconnected systems by using the hybrid system theory with memory has been considered. Agriculture picking multi-robot systems consist of multiple robots and interactions among robots, which can be regarded as a kind of multi-agent system. Distributed control of the multi-agent system is also an important issue in the study of multi-agent systems. Cooperative control tasks of multi-agent include clustering, swarming, clustering, formation, tracking, and other tasks. These collective behaviors can all be unified as consensus problems. As the core problem of distributed cooperative control, consensus means that the state values of all agents tend to be the same as time goes to infinity. The average consensus problem is a special consensus problem, which means that the final convergence value of all agents is the average of the initial value (Hu et al., 2020). In smart agriculture, average consensus plays an important role in improving the precision of agriculture. It can better realize the state consensus among agents through the initial state value of each agent, and complete cooperative agricultural tasks among agricultural multi-agents. Particularly, the average consensus of multi-agent systems also has been widely concerned in computer science, energy ecology, social economics, and other fields.

The development of distributed cooperative control has aroused the attention of many researchers in the fields of automatic control. In Olfati-Saber and Murray (2004), the Laplacian potential has been introduced for calculating the difference among agents, with which the distributed control protocol can be induced to ensure the average consensus of multi-agent systems under the undirected graph. The consensus problems have been extended to multi-agent systems whose topologies are switching (Ren and Beard, 2005), in which the consensus objective can be guaranteed if the topologies are joint connected. In the practical application, due to the limitation of speed and bandwidth of network transmission, the problem of time delay certainly exists (Sun and Wang, 2009; Chen et al., 2017; Yan and Huang, 2017). Besides, how to improve the convergence rate should be considered. In Hu et al. (2019), Zou et al. (2019), Dong et al. (2020), and Ran et al. (2020), the distributed control protocols and their convergence analyses have been investigated for finite-time consensus of multi-agent systems.

At present, continuous control protocols are used in most research, so the requirements for network communication are increased. Sampling control can not only reduce the control cost of the system by reducing information transfer redundancy but also improve the robustness of the system (Guan et al., 2012; Ding and Zheng, 2016; Park et al., 2016). Therefore, how to select the appropriate sampling time and sampling mechanism has become a concern of researchers. In Gao et al. (2009), a sampling control protocol is proposed to analyze the consensus of multi-agent systems with fixed and switched topologies, respectively. In Gao and Wang (2011), the consensus of second-order dynamical systems with time-varying topologies is studied by sampling data. The study shows that the system can be consensus by designing appropriate controller gain and the sampling period when the union graph has a spanning tree. To the best of our knowledge, there are quite limited results concerning distributed averaging problems of multi-agent systems, which are suitable for practical applications (e.g., picking multi-robot systems).

Motivated by the above discussions, we aim to reduce information transfer redundancy and control costs of smart agricultural multi-agent systems by sampling control. There are many kinds and complex communication structures in agricultural multi-agent systems. In this article, we explore the distributed averaging problems of the multi-robot systems for vegetable picking agriculture under directed communication topologies. We design a distributed control protocol by taking advantage of the nearest neighbor sampled information for the agriculture picking multi-robot system. With this protocol being used, the necessary and sufficient conditions can be provided for the average consensus objective of the agriculture picking multi-robot system. We further give how to select the sampled period. Besides, in the practical application of multi-agent systems, communication delays may exist. Especially when carrying out agricultural cooperation tasks, information needs to be transmitted between agents. Excessive communication delay will lead to the failure of agricultural multi-agent systems to achieve state consensus and thus fail to complete agricultural tasks. Therefore, we also consider the existence of time-delay among agents. We use the bilinear transformation method to develop the associated convergence analysis, which can simultaneously determine the time margin. It is shown that the agriculture picking multi-robot system can achieve the average consensus if and only if both sampled period and time delay satisfies the appropriate conditions. In addition, simulation examples are carried out to verify the theoretical results. Different from other existing sampling control studies, we carry out an average consensus analysis on picking multi-robot systems with fixed communication topology as a directed graph. The major contributions of this article include:

(i) the average consensus of agricultural multi-agent systems with fixed communication topology as a digraph is studied, and a distributed control protocol based on sampling information is proposed to ensure the average consensus of the agricultural multi-agent system.

(ii) considering the communication time-delay, another communication protocol is proposed to ensure that the agricultural multi-agent systems can achieve average consensus even with time-delay.
(iii) through the proposed distributed control protocols, the necessary and sufficient conditions for the system to achieve average consensus without time-delay and with time-delay are obtained. Because the average consensus of the system is achieved through the proposed control protocols, the relationship between the initial value and the state value is determined, and the control of the system can be realized more conveniently.

The remainder is outlined below. In Section 2, we introduce some preliminaries for graph and matrix theory. In Section 3, we introduce the problem statement for the distributed averaging of agricultural multi-agent systems for vegetable picking. Section 4 addresses the average consensus of directed the agriculture picking multi-robot system without and with communication time-delay, respectively. Simulation examples are introduced in Section 5. We give comprehensive conclusions in Section 6.

2. PRELIMINARIES FOR NOTATIONS AND GRAPH THEORY

In this section, some basic concepts related to notations and graph theory are introduced as follows.

We denote $\mathcal{I}_n = \{1, 2, \cdots, n\}$, R^n is the set of n-dimensional real numbers, $\mathrm{R}^{n\times n}$ is a set of real matrices with n-dimensional row vectors and n-dimensional column vectors, $1_n = [1, 1, \cdots, 1]^T \in \mathrm{R}^n$, and $0_n = [0, 0, \cdots, 0]^T \in \mathrm{R}^n$, and diag$\{\alpha_1, \alpha_2, \cdots, \alpha_n\}$ is a diagonal matrix in which diagonal elements are $\alpha_1, \alpha_2, \cdots, \alpha_n$, zero are the value of non-diagonal elements.

Let $\mathcal{G} = (\mathcal{V}, \mathcal{E}, A)$ represent a directed digraph, where $\mathcal{V} = \{v_1, v_2, \cdots, v_n\}$ is a node set, $\mathcal{E} \subseteq \mathcal{V} \times \mathcal{V}$ is an edge set and $A = [a_{ij}] \in \mathrm{R}^{n\times n}$ with $a_{ij} > 0 \Leftrightarrow (v_j, v_i) \in \mathcal{E}$ and $a_{ij} = 0 \Leftrightarrow (v_j, v_i) \notin \mathcal{E}$. Besides, all neighbors of v_i are denoted as $N(v_i) = \{v_j : (v_j, v_i) \in \mathcal{E}\}(i \neq j)$. An edge $(v_i, v_j) \in \mathcal{E}$ means that the data can be transmitted from the node v_i to the node v_j. Let $\mathcal{Q} = \{(v_i, v_{r_1}), (v_{r_2}, v_{r_3}), \cdots, (v_{r_n}, v_j)\}$ stand for a directed way from v_i to v_j, where $v_i, v_{r_1}, v_{r_2}, \cdots, v_{r_n}, v_j$ are different. The digraph $\mathcal{G}$ is called strongly connected on condition that $\mathcal{G}$ have at least one directed way between any twain different nodes. Let $\Delta = diag\{\sum_{j=1}^n a_{1j}, \sum_{j=1}^n a_{2j}, \cdots, \sum_{j=1}^n a_{nj}\}$ be the in-degree matrix of $\mathcal{G}$. The Laplacian matrix $L \in \mathrm{R}^{n\times n}$ of $\mathcal{G}$ is defined as $L = \Delta - A$. Benefiting from L, we can introduce the following diagonal matrix

$$W = \mathrm{diag}\{\det(L_{11}), \det(L_{22}), \cdots, \det(L_{nn})\}$$

where $L_{ii} \in \mathrm{R}^{(n-1)\times(n-1)}$ is induced from the Laplacian matrix L by deleting its ith row and ith column, det (L_{ii}) represents the determinant value of the matrix L_{ii}. The Laplacian matrix L of a strongly connected digraph $\mathcal{G}$ has a zero eigenvalue and all other eigenvalues with positive real parts. It follows from Li and Jia (2009) that if $\mathcal{G}$ is strongly connected, then $w_l = [\det(L_{11}), \det(L_{22}), \cdots, \det(L_{nn})]^T$ is the left eigenvector of L associated with the zero eigenvalue, i.e., $w_l^T L = 0_n^T$. With the help of W, we can construct a new graph $\overline{\mathcal{G}} = (\mathcal{V}, \overline{\mathcal{E}}, \overline{A})$, where the element $\overline{a}_{ij}$ of $\overline{A} = [\overline{a}_{ij}]$ satisfies

$$\overline{a}_{ij} = \frac{det(L_{ii})a_{ij} + det(L_{jj})a_{ji}}{2}, \ \forall i, j \in \mathcal{I}_n. \tag{1}$$

By taking advantage of (1), we can easily obtain

$$\overline{A} = \frac{WA + A^T W}{2}. \tag{2}$$

With (2), we can establish the relationship between L and $\overline{L}$.

Lemma 1. *For any digraph $\mathcal{G}$ and $\overline{\mathcal{G}}$, let L and $\overline{L}$ be the Laplacian matrix of $\mathcal{G}$ and $\overline{\mathcal{G}}$, respectively. The Laplacian matrices L and $\overline{L}$ satisfy*

$$\overline{L} = \frac{WL + L^T W}{2}. \tag{3}$$

proof We can directly derive this result with the help of (2) and the definition of Laplacian matrices

Lemma 2. *Consider a partitioned matrix $P = \begin{bmatrix} A & B \\ C & D \end{bmatrix} \in \mathrm{R}^{(r+s)\times(r+s)}$, where $A \in \mathrm{R}^{r\times r}$, $B \in \mathrm{R}^{r\times s}$, $C \in \mathrm{R}^{s\times r}$, and $D \in \mathrm{R}^{s\times s}$.*

1. *If A is an invertible matrix, then $|P| = |A||D - CA^{-1}B|$;*
2. *If D is an invertible matrix, then $|P| = |D||A - CD^{-1}B|$.*

In the analysis of discrete-time systems, by using a bilinear transformation, the problem of determining Schur stability of a discrete-time system can be transformed into the problem of determining Hurwitz stability of a continuous-time system. Given a polynomial with complex coefficients:

$$g(s) = \rho_n s^n + \rho_{n-1} s^{n-1} + \cdots + \rho_1 s + \rho_0 \tag{4}$$

where $\rho \in \mathrm{C}, i \in \mathcal{I}_n$. Perform a bilinear transformation $z = \frac{s+1}{s-1}$ on $g(s)$, a new polynomial is deduced

$$f(z) = (z-1)^n g(\frac{z+1}{z-1}) = \chi_0 + \chi_1 z + \cdots + \chi_n z^n \tag{5}$$

where $\chi_i = a_i + \iota b_i, a_i, b_i \in \mathrm{R}$, i = 0, 1,…, n. The Hurwitz stability of $f(z)$ implies the Schur stability of $g(s)$. Substituting $z = w\iota$ into $f(z)$, we get

$$f(\omega) = f_\omega(\omega) + \iota f_\iota(\omega) \tag{6}$$

where $f_\omega(\omega), f_\iota(\omega) \in \mathrm{R}(\omega)$, and

$$f_\omega(\omega) = a_0 - b_1\omega - a_2\omega^2 + b_3\omega^3 + a_4\omega^4 - \cdots \tag{7}$$

$$f_\iota(\omega) = b_0 + a_1\omega - b_2\omega^2 - a_3\omega^3 + b_4\omega^4 - \cdots \tag{8}$$

$f_\omega(\omega)$ and $f_\iota(\omega)$ constituent interlaced polynomial, to determine whether $f(z)$ Hurwitz stable, the Hermite-Biehler theorem is given as follows

Lemma 3. *(Ogata, 1995) The polynomial $f(z)$ is Hurwitz stability if and only if the related pair $f_\omega(\omega)$, $f_\imath(\omega)$ is interlaced, and $f_\omega(0)f_\imath'(0) - f_\omega'(0)f_\imath(0) > 0$.*

3. PROBLEM STATEMENTS

In the picking process of vegetables, such as cucumbers, a device with multiple mechanical arms is used to pick the vegetables. In the actual picking process, multiple mechanical arms pick vegetables at the same time and then put vegetables into the picking robot which runs on a specific track. To accurately collect vegetables picked by mechanical arms, the corresponding collection robots should achieve position states consensus, to better pick and collect vegetables. To solve the problem, a consensus analysis of the robot's position state is needed. In this article, we treat each robot as an intelligent agent and all agents constitute an agriculture multi-agent system. To analyze the consensus of the agriculture multi-agent system, we use a digraph $\mathcal{G} = (\mathcal{V}, \mathcal{E}, A)$ of the agricultural multi-agent system to denote the communication topology of the agriculture picking multi-robot system, in which the set of all agents can be described by $\mathcal{V}$, and the relationships among agents can be represented by $\mathcal{E}$ and A. Let $x_i \in \mathrm{R}^n$ be the position state of agent v_i and $x(t) = [x_1(t), x_2(t), \cdots, x_n(t)]^T \in \mathrm{R}^n$ denote the state vector. Every agent has the following dynamics

$$\dot{x}_i(t) = u_i(t),\ \forall i \in \mathcal{I}_n \tag{9}$$

where u_i is the control protocol to be designed.

Generally, the decision value of the agriculture multi-agent system not only depends on the topological structure but also on the initial states. However, the average-consensus problems only rely on the initial states and have no relation to the topological structure. That is to say, for random initial states $x_i(0), \forall i \in \mathcal{I}_n$, the average consensus of the system (9) can be reached if

$$\lim_{t\to\infty} x_i(t) = \frac{1}{n}\sum_{i=1}^{n} x_i(0), \forall i \in \mathcal{I}_n. \tag{10}$$

The agricultural multi-agent system (9) can reach average consensus means that we can infer the final position of agents from the initial position, so the control difficulty is reduced, and the controllability of the agriculture multi-agent system (9) is improved.

Since continuous control will increase the communication burden of agricultural multi-agent systems, in order to prevent information redundancy and reduce the cost of systems, we use sampling data to complete the distributed control of the agricultural multi-agent system. The sampling control can improve the robustness of the picking robot system. In what follows, the purpose of this article is to design a distributed control protocol so that the agricultural multi-agent system (9) under the strongly connected digraph $\mathcal{G}$ accomplishes the average consensus objective *via* sampled information. Besides, when considering the communication time-delay, we further explore how to develop the time margin of the communication time-delay.

4. MAIN RESULTS

In this section, we investigate the average consensus problems of the agricultural multi-agent system (9) whose communication topologies are directed. Besides, information needs to be transmitted between agents, and the excessive communication time-delay will cause the oscillation or divergence of agricultural multi-agent systems so that the robots in agriculture cannot achieve position state consensus, which means that the robots cannot accurately load the picked vegetables. Thus, we further explore the average consensus problems of agricultural multi-agent systems when there exist communication time-delays among agents.

4.1. Distributed Control Protocol Without Time-Delay

In this subsection, to reduce the communication cost of smart agricultural multi-agent systems, we aim to solve the average consensus problems of the agriculture multi-agent system by taking advantage of sampled data. Toward this end, we introduce a distributed control protocol by employing the sampled data as follows:

$$u_i(t) = \sum_{j\in N(i)} \overline{a}_{ij}(x_j(kp) - x_i(kp)), t \in [kp, kp + p), \quad \forall i, j \in \mathcal{I}_n,\ k = 0, 1, 2, \cdots \tag{11}$$

where p is the sampled period. Based on $\overline{L}$, we can rewrite (9) and (11) as a compact form

$$x(kp + p) = \overline{\psi}x(kp),\ k = 0, 1, 2, \cdots \tag{12}$$

where $\overline{\psi} = I - p\overline{L}$.

In the following, we explore the convergence problems of the system (12). We propose a tree transformation for the system (12). We first introduce a series of states as follows:

$$\begin{aligned} y_1(kp) &= x_1(kp) \\ y_2(kp) &= x_1(kp) - x_2(kp) \\ y_3(kp) &= x_1(kp) - x_3(kp) \\ &\vdots \\ y_n(kp) &= x_1(kp) - x_n(kp). \end{aligned} \tag{13}$$

Denote

$$Q = \begin{bmatrix} 1 & 0 & 0 & \cdots & 0 \\ 1 & -1 & 0 & \cdots & 0 \\ 1 & 0 & -1 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & 0 & 0 & \cdots & -1 \end{bmatrix} = \begin{bmatrix} C \in \mathrm{R}^{1\times n} \\ E \in \mathrm{R}^{(n-1)\times n} \end{bmatrix}$$

and the inverse matrix of Q is given by

$$Q^{-1} = \begin{bmatrix} 1 & 0 & 0 & \cdots & 0 \\ 1 & -1 & 0 & \cdots & 0 \\ 1 & 0 & -1 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & 0 & 0 & \cdots & -1 \end{bmatrix} = \begin{bmatrix} w_r \in \mathrm{R}^{n\times 1} & F \in \mathrm{R}^{n\times(n-1)} \end{bmatrix}.$$

With the help of Q, the states $y_1(kp)$, $y_2(kp)$, $\cdots$, $y_n(kp)$ are defined by

$$y(kp) \triangleq \begin{bmatrix} y_1(kp) \\ y_2(kp) \\ \vdots \\ y_n(kp) \end{bmatrix} = Qx(kp).$$

Substituting $y(kp+p) = Qx(kp+p)$ and $y(kp) = Qx(kp)$ into the system (12) leads to

$$y(kp+p) = (I - pH)y(kp) \tag{14}$$

where $H = Q\overline{L}Q^{-1} = \begin{bmatrix} 0 & C\overline{L}F \\ 0_{n-1} & E\overline{L}F \end{bmatrix}$. We denote $\hat{y}(kp) = [y_2(kp), y_3(kp), \cdots, y_n(kp)]^T$ and the system (14) can be divided into two subsystems:

$$y_1(kp+p) = y_1(kp) - pC\overline{L}F\hat{y}(kp) \tag{15}$$

and

$$\hat{y}(kp+p) = (I - pE\overline{L}F)\hat{y}(kp). \tag{16}$$

From (16), we can easily see that the reduced system (16) achieving stability implies the consensus of the system (12). Hence, the consensus problem of (12) turns into the asymptotic stability problem of a reduced-order system (16).

With protocol (11) being employed, the average consensus results can be obtained in the following theorem.

Theorem 1. *For the system (9) whose communication topology is the strongly connected digraph $\mathcal{G}$, let the distributed control protocol (13) be used. Then, the system (9) can achieve the average consensus objective if and only if the following condition holds.*

$$0 < p < \min_{\lambda_i \neq 0} \frac{2}{\lambda_i}, \ \forall i \in \mathcal{I}_n. \tag{17}$$

Proof. The digraph $\mathcal{G}$ is strongly connected, its Laplacian matrix L has a zero eigenvalue and $n-1$ non-zero eigenvalue with positive real parts. It follows from Lemma1 that $\overline{L}$ is a symmetric matrix. The eigenvalues of $\overline{L}$ contain a zero eigenvalue and $n-1$ positive real numbers. Based on our defined H, we realize that the eigenvalues of $E\overline{L}F$ are positive real numbers. To ensure the reduced-order system (16) is stable, the condition $\rho(I - pE\overline{L}F) < 1$. Next, we target at exploring how to ensure $\rho(I - pE\overline{L}F) < 1$ by picking up the sampled period p.

We introduce an inverse matrix T such that

$$T^{-1}E\overline{L}FT = \Lambda = \begin{bmatrix} \lambda_2 & * & & \\ & \lambda_3 & \ddots & \\ & & \ddots & * \\ & & & \lambda_n \end{bmatrix} \tag{18}$$

where $\lambda_2, \lambda_3, \cdots, \lambda_n$ are the nonzero eigenvalues of $\overline{L}$ and the elements $*$ maybe 0 or 1. Then, employing $\tilde{y}(kp+p) = P^{-1}\hat{y}(kp+p)$, $\tilde{y}(kp) = P^{-1}\hat{y}(kp)$, we can convert the reduced-order system (16) into

$$\tilde{y}(kp+p) = (I - p\Lambda)\tilde{y}(kp). \tag{19}$$

The stability of the systems (16) and (19) are equivalent. We can further induce

$$\rho(I - p\Lambda) < 1 \Leftrightarrow |1 - p\lambda_i| < 1 \text{ for all eigenvalues of } \overline{L}. \tag{20}$$

Therefore, $\rho(I - p\Lambda) < 1$ holds if and only if p meets the condition (17) holds. Based on the condition (17), the reduced system (16) can reach stability, which denotes that the system (9) is able to achieve the consensus.

Next, we calculate the convergence value of the dynamic system (9). Let w_r and w_l denote the right eigenvector and left eigenvector of $\overline{L}$ associated with its eigenvalue 0, respectively, which satisfy $w_l^T w_r = 1$. Correspondingly, we can easily obtain that w_r and w_l are also the right eigenvector and left eigenvector of $\overline{\psi} = I - p\overline{L}$ associated with the eigenvalue 1. Since the digraph $\mathcal{G}$ is strongly connected, we can develop that $\overline{\mathcal{G}}$ is undirected and connected. Without loss of generality, we select w_r and w_l that satisfy $w_r = w_l = \frac{1}{\sqrt{n}}1_n$. With (9), we can deduce

$$\begin{aligned} x(kp+p) &= \overline{\psi}x(kp) \\ &= \overline{\psi}^{k+1}x(0). \end{aligned}$$

Because $\overline{\psi}$ has an eigenvalue 1 and $n-1$ eigenvalues whose module is less than 1. Hence, we can develop

$$\begin{aligned} \lim_{k\to\infty} x(kp) &= \lim_{k\to\infty} \overline{\psi}^k x(0) \\ &= w_r w_l^T x(0) = \frac{1}{n}\sum_{i=1}^{n} x_i(0) \end{aligned}$$

which implies the system (9) can achieve the average consensus objective *via* sampled control. The sampled control can be used to reduce the communication cost of the agricultural multi-agent system (9). We complete this proof.

4.2. Distributed Control Problems With Time-Delay

It is shown that under the proposed protocol without time-delay, the states of all agents can be guaranteed to reach a consensus whose value is the averaging of the initial states of all agents. However, in the practical application of agriculture multi-agent

systems, it may suffer from the effect of communication time-delay. Especially in the completion of agricultural cooperative tasks, information needs to be transmitted between agents, and the excessive communication delay will cause the oscillation or divergence of multi-agent systems, so the time-delay problem needs to be considered. When considering the existing time-delay τ which is less than 1 sampling period, the distributed control protocol is constructed by

$$u_i(t) = \begin{cases} \sum_{j\in N(i)} \bar{a}_{ij}(x_j(kp-p) - x_i(kp-p)), t \in [kp, kp+\tau) \\ \sum_{j\in N(i)} \bar{a}_{ij}(x_j(kp) - x_i(kp)), t \in [kp+\tau, kp+p) \end{cases}$$
$$\forall i, j \in \mathcal{I}_n,\ k = 0, 1, 2, \cdots,\ 0 < \tau < p. \tag{21}$$

We rewrite (9) and (21) as follows

$$\begin{bmatrix} x(kp+p) \\ x(kp) \end{bmatrix} = \phi \begin{bmatrix} x(kp) \\ x(kp-p) \end{bmatrix}, k = 0, 1, 2, \cdots \tag{22}$$

where

$$\phi = \begin{bmatrix} I_n - (p-\tau)\bar{L} & -\tau\bar{L} \\ I_n & 0 \end{bmatrix}. \tag{23}$$

We can deduce from (22) and (23) that

$$x(kp+p) = [I_n - (p-\tau)\bar{L}]x(kp) - \tau\bar{L}x(kp-p), \quad k = 0, 1, 2, \cdots \tag{24}$$

In the following, we also use the tree transformation for the system (22), the consensus problem of (22) turns into the asymptotic stability problem of a reduced-order system.

Substituting $y(kp+p) = Qx(kp+p)$, $y(kp) = Qx(kp)$, and $y(kp-p) = Qx(kp-p)$ into the system (22) leads to

$$\begin{bmatrix} y(kp+p) \\ y(kp) \end{bmatrix} = \begin{bmatrix} I_n - (p-\tau)H & -\tau H \\ I_n & 0_{n\times n} \end{bmatrix} \begin{bmatrix} y(kp) \\ y(kp-p) \end{bmatrix} \tag{25}$$

where $H = Q\bar{L}Q^{-1} = \begin{bmatrix} 0 & C\bar{L}F \\ 0_{n-1} & E\bar{L}F \end{bmatrix}$.

We denote

$$\hat{y}(kp) = [y_2(kp), y_3(kp), \cdots, y_n(kp)]^T,$$
$$\hat{y}(kp-p) = [y_2(kp-p), y_3(kp-p), \cdots, y_n(kp-p)]^T$$

and the system (25) can be divided into two subsystems:

$$\begin{bmatrix} y_1(kp+p) \\ y_1(kp) \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} y_1(kp) \\ y_1(kp-p) \end{bmatrix} - \begin{bmatrix} (p-\tau)C\bar{L}F & \tau C\bar{L}F \\ 0 & 0 \end{bmatrix} \begin{bmatrix} \hat{y}(kp) \\ \hat{y}(kp-p) \end{bmatrix} \tag{26}$$

and

$$\begin{bmatrix} \hat{y}(kp+p) \\ \hat{y}(kp) \end{bmatrix} = \begin{bmatrix} I_{n-1} - (p-\tau)E\bar{L}F & -\tau E\bar{L}F \\ I_{n-1} & 0_{(n-1)\times(n-1)} \end{bmatrix} \begin{bmatrix} \hat{y}(kp+p) \\ \hat{y}(kp-p) \end{bmatrix}. \tag{27}$$

From (27), we can easily see that the reduced system (27) achieving stability implies the consensus of the system (9). With the protocol (21) being employed, the average consensus results can be obtained in the theorem below.

Theorem 2. *For the system (9) whose communication topology is the strongly connected digraph $\mathcal{G}$, let the distributed control protocol (21) be used. Then, the system (9) can achieve the average consensus objective if and only if the following condition holds.*

$$\tau < \min_{\lambda_i\neq 0}\frac{1}{\lambda_i}, \quad 0 < p < \min_{\lambda_i\neq 0}\frac{2}{\lambda_i} + 2\tau,\ \forall i \in \mathcal{I}_n. \tag{28}$$

Proof. Since the digraph $\mathcal{G}$ is strongly connected, its Laplacian matrix L has a zero eigenvalue and $n-1$ non-zero eigenvalue with positive real parts. It follows from Lemma1 that $\bar{L}$ is a symmetric matrix.

We introduce an inverse matrix T such that

$$T^{-1}E\bar{L}FT = \Lambda = \begin{bmatrix} \lambda_2 & * & & \\ & \lambda_3 & \ddots & \\ & & \ddots & * \\ & & & \lambda_n \end{bmatrix} \tag{29}$$

where $\lambda_2, \lambda_3, \cdots, \lambda_n$ are the nonzero eigenvalues of $\bar{L}$ and the elements $*$ may be 0 or 1. Then, employing $\tilde{y}(kp+p) = T^{-1}\hat{y}(kp+p)$, $\tilde{y}(kp) = T^{-1}\hat{y}(kp)$, and $\tilde{y}(kp-p) = T^{-1}\hat{y}(kp-p)$ we can convert the reduced-order system (27) into

$$\begin{bmatrix} \tilde{y}(kp+p) \\ \tilde{y}(kp) \end{bmatrix} = \zeta \begin{bmatrix} \tilde{y}(kp+p) \\ \tilde{y}(kp-p) \end{bmatrix} \tag{30}$$

where

$$\zeta = \begin{bmatrix} I_{n-1} - (p-\tau)\Lambda & -\tau\Lambda \\ I_{n-1} & 0_{(n-1)\times(n-1)} \end{bmatrix}. \tag{31}$$

The stability of the systems (27) and (30) are equivalent. The characteristic polynomial of ζ is given by

$$det(sI_{2n-2} - \zeta) = \begin{vmatrix} sI_{n-1} - [I_{n-1} - (p-\tau)\Lambda] & \tau\Lambda \\ -I_{n-1} & sI_{n-1} \end{vmatrix}. \tag{32}$$

TABLE 1 | Comparison of conditions for achieving average consensus in distributed control protocols.

Number	Distributed control protocols	Conditions
1	$u_i(t) = \sum_{j\in N(i)} \bar{a}_{ij}(x_j(kp) - x_i(kp)), t \in [kp, kp+p)$	$0 < p < \min\frac{2}{\lambda_i}$
2	$u_i(t) = \begin{cases} \sum_{j\in N(i)} \bar{a}_{ij}(x_j(kp-p) - x_i(kp-p)), t \in [kp, kp+\tau) \\ \sum_{j\in N(i)} \bar{a}_{ij}(x_j(kp) - x_i(kp)), t \in [kp+\tau, kp+p) \end{cases}$	$\tau < \min\frac{1}{\lambda_j}$ $0 < p < \min\frac{2}{\lambda_i} + 2\tau$

From (32), we can easily obtain that ξ is a partitioned matrix. It follows from Lemma 2 that

$$\begin{aligned}
|\xi| &= |sI_{n-1}|\left|sI_{n-1} - [I_{n-1} - (p-\tau)\Lambda] + I_{n-1}(sI_{n-1})^{-1}\tau\Lambda\right| \\
&= \begin{vmatrix} s & & \\ & \ddots & \\ & & s \end{vmatrix} \begin{vmatrix} s - [1-(p-\tau)\lambda_2] + \frac{1}{s}\tau\lambda_2 & & \\ & \ddots & \\ & & s - [1-(p-\tau)\lambda_n + \frac{1}{s}\tau\lambda_n] \end{vmatrix} \\
&= \prod_{i=2}^{n}[s^2 - (1 - p\lambda_i + \tau\lambda_i)s + \tau\lambda_i] \\
&= \prod_{i=2}^{n} g_i(s).
\end{aligned} \tag{33}$$

Then, by applying bilinear transformation $s = \dfrac{z+1}{z-1}$, we have

$$f_i(z) = p\lambda_i z^2 + 2(1 - \tau\lambda_i)z + (p - 2\tau)\lambda_i + 2 \tag{34}$$

We can prove the polynomial $g_i(s)$ is Schur stable by making sure that the polynomial (34) is Hurwitz stable. Let $z = \omega\iota$, we can further deduce

$$f_i(\omega) = -p\lambda_i\omega^2 + 2(1 - \tau\lambda_i)\omega\iota + (p - 2\tau)\lambda_i + 2. \tag{35}$$

The real part and imaginary part of (35) are given by

$$f_\omega(\omega) = -p\lambda_i\omega^2 + (p - 2\tau)\lambda_i + 2 \tag{36}$$

and

$$f_\iota(\omega) = 2(1 - \tau\lambda_i)\omega\iota \tag{37}$$

The polynomial (35) is Hurwitz stable if and only if the following conditions hold.

C1) $f_\omega(\omega) = 0$ has two distinct roots $\gamma_1 < \gamma_2$.
C2) The real root γ_3 of $f_\iota = 0$ satisfies $\gamma_1 < \gamma_3 < \gamma_2$.
C3) $f_\omega(0)f_\iota^{'}(0) - f_\omega^{'}(0)f_\iota(0) > 0$.

The condition C1) can be guaranteed by

$$\Delta_{f_\omega} = 4(p\lambda_i)[2 - (p - 2\tau)\lambda_i] > 0. \tag{38}$$

Noticing $p > 0$ and $\lambda_i (i = 2, 3, \ldots, N)$. Based on (38), we further induce

$$0 < p < \min_{\lambda_i \neq 0}\frac{2}{\lambda_i} + 2\tau,\ \forall i \in \mathcal{I}_n. \tag{39}$$

If $\Delta_{f_\omega} > 0$, we can calculate two roots γ_1 and γ_2 of $f_\omega(\omega) = 0$ as follows:

$$\gamma_1 = -\frac{\sqrt{\Delta_{f_\omega}}}{2p\lambda_i}, \gamma_2 = \frac{\sqrt{\Delta_{f_\omega}}}{2p\lambda_i}.$$

Based on $f_\iota = 0$, we can get its root $\gamma_3 = 0$. The condition C2) is naturally satisfied. It is mainly because that $\gamma_1 < 0$ and $\gamma_2 > 0$ hold. Motivated by the condition C3), we can deduce

$$[2 - (h - 2\tau)\lambda_i](1 - \tau\lambda_i) > 0. \tag{40}$$

With (39) and (40), we can develop that the system (35) is Hurwitz stable if and only if p and τ meet (28). Therefore, we can develop that the reduced system (27) achieves asymptotic stability. It denotes that the system (22) can achieve the consensus objective.

Next, we calculate the convergence value of the system (22). One of the eigenvalues of $\bar{L}$ is 0, Correspondingly, we can infer that 1 are the eigenvalues of ϕ and the module value of other eigenvalues is less than 1. Because $\bar{\mathcal{G}}$ is undirected and connected, we can pick up the right eigenvector and the left eigenvector of the matrix ϕ corresponding to the eigenvalue 1 as follows:

$$w_r = w_l = \frac{1}{\sqrt{n}}1_n$$

which satisfy $w_l^T w_r = 1$. According to (24), we have

$$\begin{bmatrix} x(kp+p) \\ x(kp) \end{bmatrix} = \phi \begin{bmatrix} x(kp) \\ x(kp-p) \end{bmatrix} = \phi^k \begin{bmatrix} x(0) \\ x(0) \end{bmatrix}.$$

Because ϕ all the other eigenvalues are in the unit circle. Hence, we can develop

$$\lim_{k\to\infty} \begin{bmatrix} x(kp+p) \\ x(kp) \end{bmatrix} = \begin{bmatrix} w_r w_l^T & 0 \\ w_r w_l^T & 0 \end{bmatrix} \begin{bmatrix} x(0) \\ x(0) \end{bmatrix} = \begin{bmatrix} \frac{1}{n}\sum_{i=1}^{n} x_i(0) \\ \frac{1}{n}\sum_{i=1}^{n} x_i(0) \end{bmatrix}$$

which implies that the system (9) can achieve the average consensus even if the sampled instance p and the communication time-delay τ satisfy the condition (28). This proof is complete.

4.3. Analysis and Comparison of Protocols

We compare and analyze the condition of achieving the average consensus of the agricultural multi-agent system without time-delay and with time-delay. The protocols in different cases and the conditions for achieving average consensus in agricultural multi-agent systems are given in **Table 1**, where $\lambda_i \neq 0$ and $\forall i, j \in \mathcal{I}_n$, $k = 0, 1, 2, \cdots, n, 0 < \tau < p$. As shown in **Table 1**, number 1 is the distributed control protocol in the case of no time-delay, it can be found that the upper limit of the sampling period depends on the eigenvalue of $\bar{L}$ through the control protocol (11) proposed by Theorem 1. For the case with time delay is number 2, by applying control protocol (21) and observation Theorem 2, we can find that the upper limit of

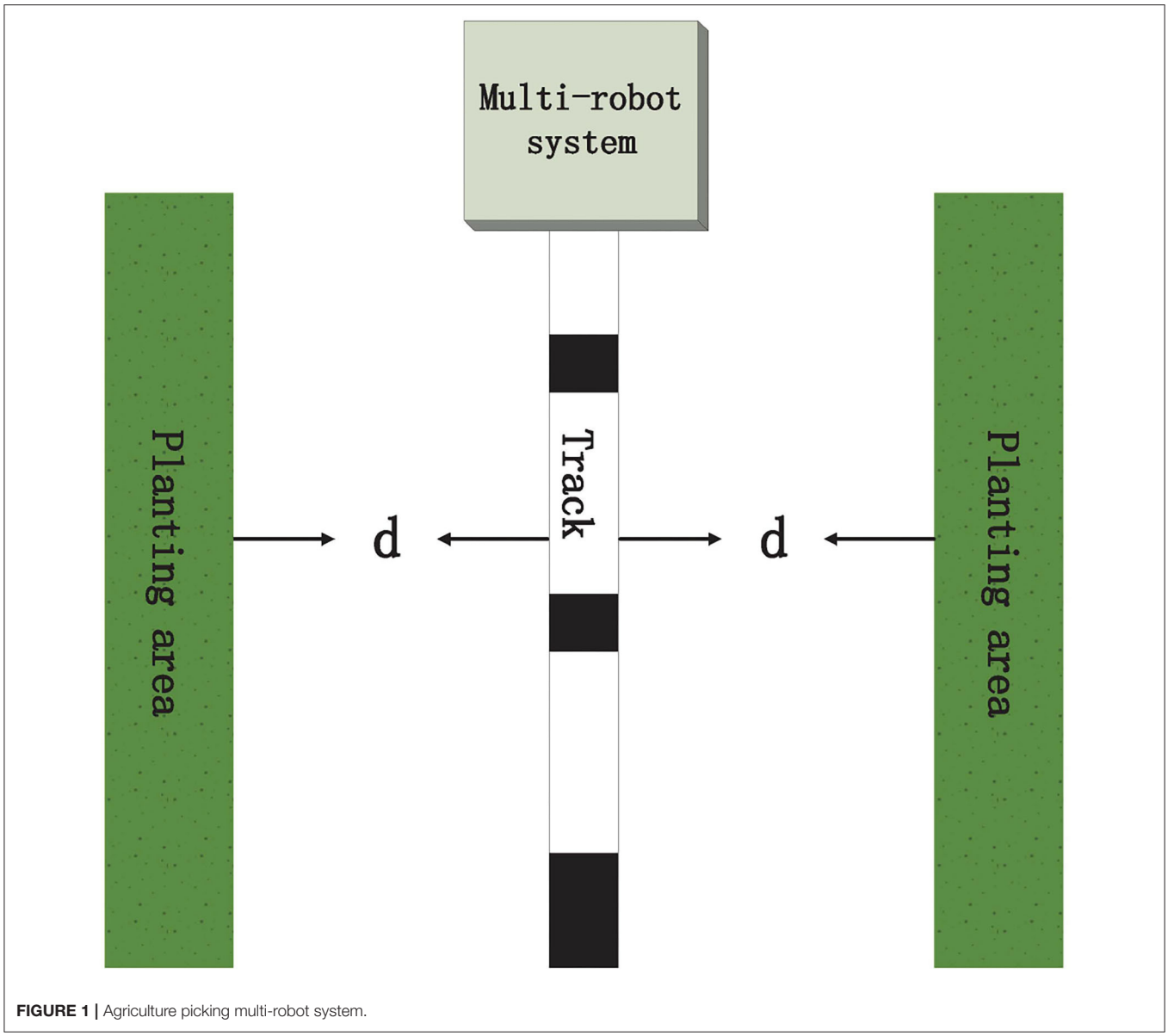


FIGURE 1 | Agriculture picking multi-robot system.

time-delay τ depends on the eigenvalue of $\overline{L}$, and the sampling period is not only related to the eigenvalue of $\overline{L}$ but also related to the value of τ. This means that on the premise that τ meets the value condition, the upper limit of sampling period p can be further calculated, and to obtain the corresponding value range of sampling period p under different time-delay τ values.

5. SIMULATIONS

This section will introduce two simulations to illustrate the correctness of our developed theoretical results. Here, we consider agriculture picking a multi-robot system including six robots in **Figure 1**, in which robots can pick the vegetables along the track. From **Figure 1**, we can see that the mechanical arm length of multi-robots needs to satisfy the distance between multi-robots and two plant areas for the purpose of conveniently gathering the vegetables. Hence, we should select a suitable orbital position to meet the distance d. Without loss of generality, we assume that the position state of the orbit is 5 such that the distance between orbit and planting area is d. Motivated by distributed averaging, we can pick up the initial states of six robots as follows:

$$x(0) = [2, 6, 7, 3, 4, 8]^T.$$

Example 1. The communication topology of multi-robots can be described in **Figure 2**. We can easily see from **Figure 2** that the digraph $\mathcal{G}_1$ is strongly connected and weight unbalanced, in which only the weight 1 is considered. We can get the

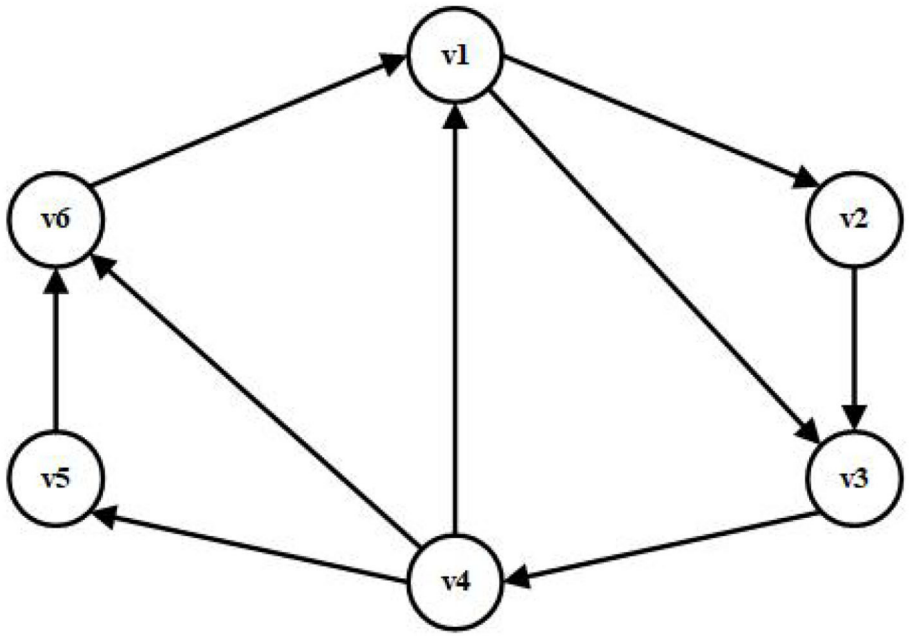


FIGURE 2 | Strong connected digraph $\mathcal{G}_1$.

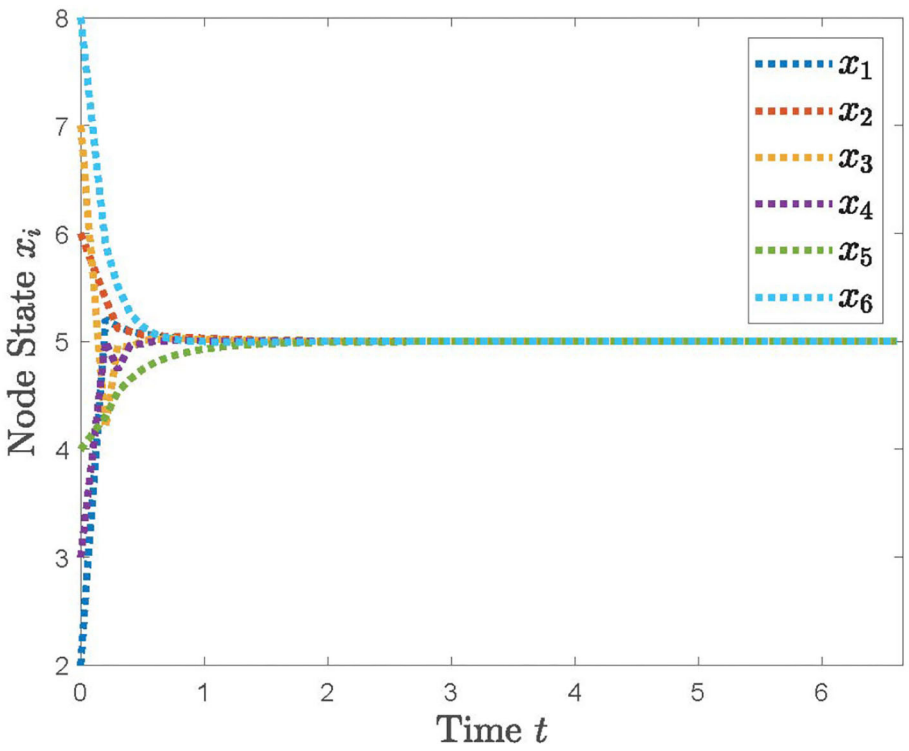


FIGURE 3 | The simulation result with $p = 0.1$.

Laplacian matrix L of $\mathcal{G}_1$. By Lemma 1, the specific values of the matrix $\overline{L}$ are computed, and the eigenvalues of $\overline{L}$ are further obtained. With the help of Theorem 1, then we can compute the upper bound value of p is 0.16198. Hence, the system (12) can reach the consensus objective if and only if $p < 0.16198$ holds.

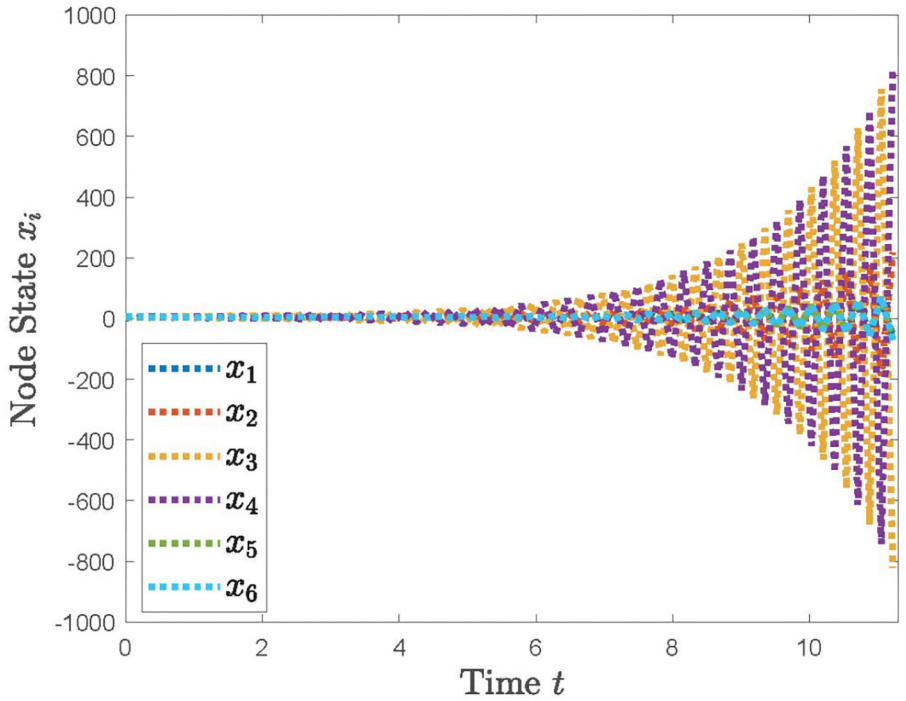


FIGURE 4 | The simulation result with $p = 0.17$.

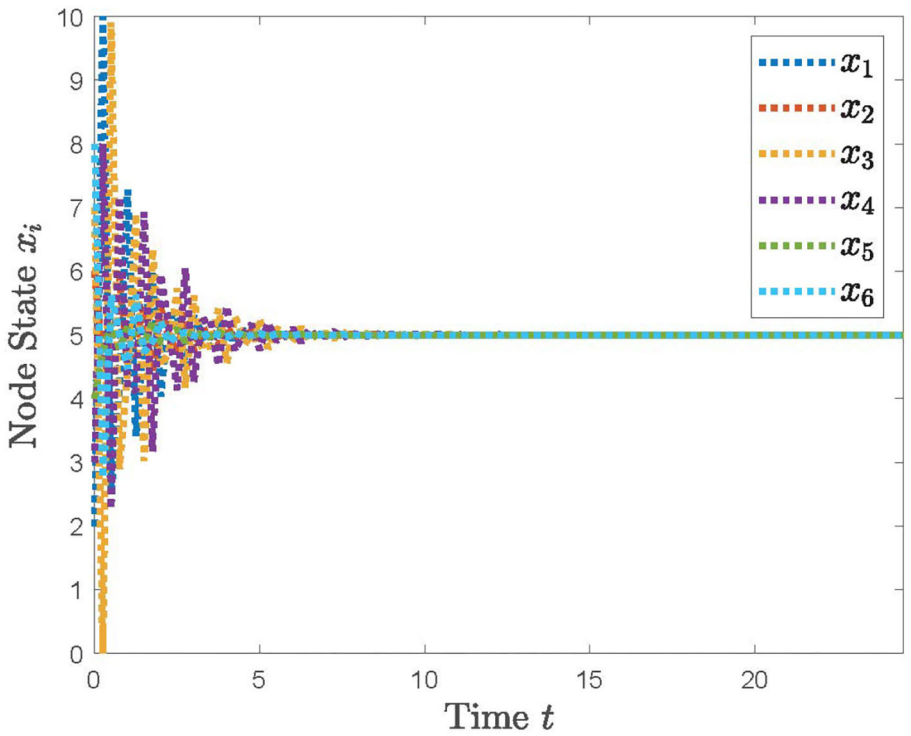


FIGURE 5 | The simulation result with $\tau = 0.06, p = 0.25$.

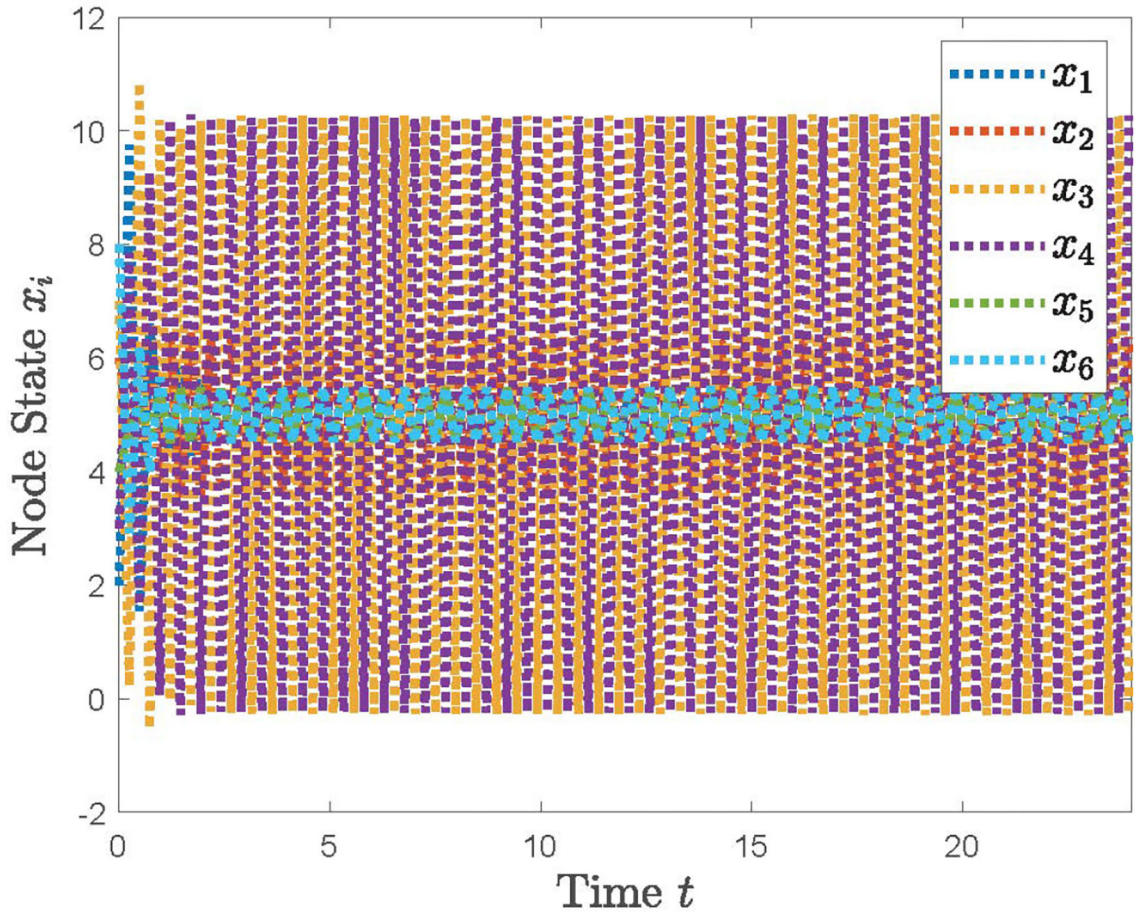


FIGURE 6 | The simulation result with $\tau = 0.04, p = 0.24198$.

With the distributed control protocol (11) being employed, the state evolution of the system (12) with the sampled period $p = 0.1$ is plotted in **Figure 3** and with the sampled period $p = 0.17$ in **Figure 4**. It is obvious from **Figures 3**, **4** that the system (12) can achieve the average consensus objective with value 5 when the sampled instant is $p = 0.1$ and the system (12) is divergent when the sampled instant is $p = 0.17$, which coincide with the results of Theorem 1. Therefore, our designed protocol can ensure the multi-robots reach the arbitrary expected position by selecting the initial states of the robots.

Example 2. When considering the communication time-delay that is inevitable in communication among agents, we can calculate that the time margin τ satisfies $\tau_{max} = 0.081$ and $p < \frac{2}{\lambda_{max}} + 2\tau$ based on (28). The state evolution of the system (22) with $\tau = 0.06, p = 0.25$, $\tau = 0.04, p = 0.24198$, and $\tau = 0.081, p = 0.3$ are plotted in **Figures 5**–**7**, respectively. From **Figures 5**–**7**, we can see that the system (12) can accomplish the average consensus objective if $\tau_{max} = 0.081$ and $p < \frac{2}{\lambda_{max}} + 2\tau$ hold and diverge, otherwise. Hence, the effectiveness of Theorem 2 can be verified.

6. CONCLUSION

To prevent information transfer redundancy and reduce the cost of the agricultural multi-agent systems, this article has investigated the distributed averaging problems of directed agriculture picking multi-robot systems with sampling control. We have designed the distributed control protocols with and without time-delay by neighbor information to accomplish the average consensus objective, respectively. It is shown that the necessary and sufficient conditions have been provided for the average consensus, and the robots of the agriculture picking multi-robot system can achieve the purpose of consensus position state by setting the initial value. Two simulation examples have been introduced to demonstrate the effectiveness of our derived results. The distributed control protocols designed by sampling control can achieve the average consensus, which means that the agents of smart agricultural multi-robot systems can achieve state consensus, and the robots of agricultural multi-robot systems can accurately collect vegetables picked by mechanical arms. The protocols we designed also can reduce the information redundancy and control costs of smart agricultural multi-agent systems by sampling control, even agricultural

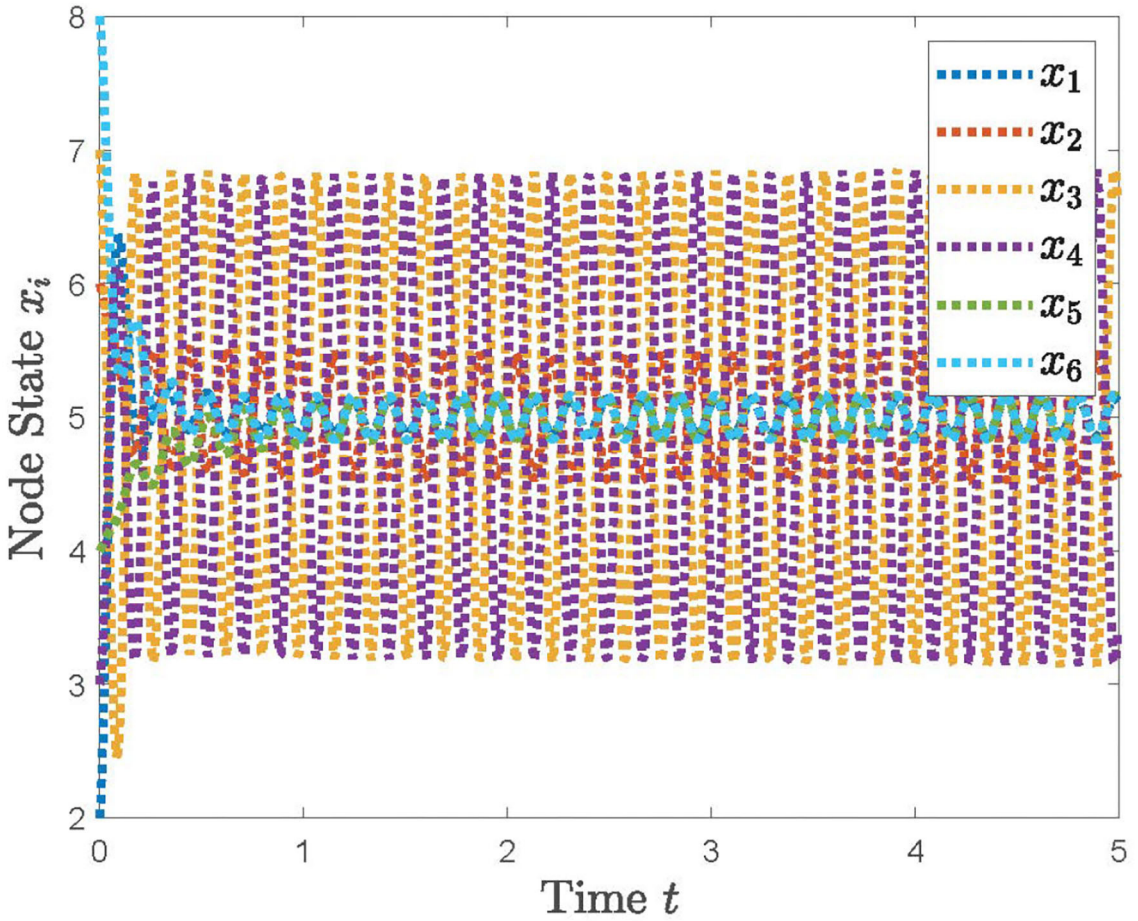


FIGURE 7 | The simulation result with $\tau = 0.081, p = 0.3$.

multi-agent systems with time-delay, and provide a feasible method to improve the mechanization level of smart agriculture. At present, the proposed protocols are only applicable to the structure proposed, and only apply to the case where the sampling period is fixed, but we have further studied whether it is still applicable in other environments. In the future, we attempt to change the problem of the fixed sampling period into the study of the variable sampling period, so that the multi-robot systems can achieve average consistency under more working conditions and further promote the smooth completion of cooperative tasks among robots.

DATA AVAILABILITY STATEMENT

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

AUTHOR CONTRIBUTIONS

FM: funding acquisition and project administration. HY: writing—original draft. MD: conceptualization, methodology, and writing—reviewing and editing. PJ: supervision. XS: software. All the authors have read and agreed to the published version of the manuscript.

FUNDING

This study was supported in part by the National Natural Science Foundation of China under Grants 62103210 and 61903207, in part by the Industry-university-research Collaborative Innovation Fund project of Qilu University of Technology (Shandong Academy of Sciences) under Grant 2020-CXY26, and in part by the Key Technology Research and Development Program of Shandong (Major Scientific and Technological Innovation Project) under Grants 2019JZZY010731 and 2020CXGC010901.

ACKNOWLEDGMENTS

The authors would like to thank the Associate Editor and two reviewers for their insightful comments and suggestions which greatly improved the quality and presentation of this manuscript.

REFERENCES

Alsamhi, S. H., Lee, Brian, Guizani, M., Kumar, N., Qiao, Y., and Liu X. (2021). Blockchain for decentralized multi-drone to combat COVID-19 and future pandemics: Framework and proposed solutions. *Trans. Emerg. Telecommun. Technol.* 32, e4255. doi: 10.1002/ett.4255

Alsamhi, S. H., Ma, O., and Ansari, M. S. (2019b). Survey on artificial intelligence based techniques for emerging robotic communication. *Telecommun. Syst.* 72, 483–503. doi: 10.1007/s11235-019-00561-z

Alsamhi, S. H., Ma, O., and Ansari, M. S. (2020). Convergence of machine learning and robotics communication in collaborative assembly: Mobility, connectivity and future perspectives. *J. Intell. Robot. Syst.* 98, 541–56. doi: 10.1007/s10846-019-01079-x

Alsamhi, S. H., Ma, O., Ansari, M. S., Gupta, S. K. (2019a). Collaboration of drone and internet of public safety things in smart cities: An overview of qos and network performance optimization. *Drones* 3, 13. doi: 10.3390/drones3010013

Brondino, L., Borra, D., Giuggioli, N., and Massaglia, S. (2021). Mechanized blueberry harvesting: preliminary results in the Italian context. *Agriculture* 11:1197. doi: 10.3390/agriculture11121197

Chen, C. L.P., Wen, G., Liu, Y., and Wang, F. (2017). Adaptive consensus control for a class of nonlinear multiagent time-delay systems using neural networks. *IEEE Trans. Neural Netw. Learn. Syst.* 25, 1217–1226. doi: 10.1109/TNNLS.2014.2302477

Chen, J., and Yang, A. (2019). Intelligent agriculture and its key technologies based on internet of things architecture. *IEEE Access* 7, 77134–77141. doi: 10.1109/ACCESS.2019.2921391

Ding, L., and Zheng, W. (2016). Consensus tracking in heterogeneous nonlinear multi-agent networks with asynchronous sampled-data communication. *Syst. Control Lett.* 96, 151–157. doi: 10.1016/j.sysconle.2016.08.001

Dong, G., Li, H., Ma, H., and Lu, R. (2020). Finite-time consensus tracking neural network ftc of multi-agent systems. *IEEE Trans. Neural Netw. Learn. Syst.* 32, 653–662. doi: 10.1109/TNNLS.2020.2978898

Fuentes, A., Yoon, S., Kim, T., and Park, D. (2021). Open set self and across domain adaptation for tomato disease recognition with deep learning techniques. *Front. Plant Sci.* 12:758027. doi: 10.3389/fpls.2021.758027

Gao, Y., and Wang, L. (2011). Sampled-data based consensus of continuous-time multi-agent systems with time-varying topology. *IEEE Trans. Autom. Control* 56, 1226–1231. doi: 10.1109/TAC.2011.2112472

Gao, Y., Wang, L., Xie, G., and Wu, B. (2009). Consensus of multi-agent systems based on sampled-data control consensus of multi-agent systems based on sampled-data control. *Int. J. Control* 82, 2193–2205. doi: 10.1080/00207170902948035

Guan, Z., Liu, Z., Feng, G., and Jian, M. (2012). Impulsive consensus algorithms for second-order multi-agent networks with sampled information. *Automatica* 48, 1397–1404. doi: 10.1016/j.automatica.2012.05.005

Horng, G., Liu, M., and Chen, C. (2019). The smart image recognition mechanism for crop harvesting system in intelligent agriculture. *IEEE Sensors J.* 20, 2766–2781. doi: 10.1109/JSEN.2019.2954287

Hu, B., Guan, Z., and Fu, M. (2019). Distributed event-driven control for finite-time consensus. *Automatica* 103, 88–95. doi: 10.1016/j.automatica.2019.01.026

Hu, X., Zhang, Z., C., L., and Zhang, Q. (2020). Convergence of machine learning and robotics communication in collaborative assembly: mobility, connectivity and future perspectives. *J. Franklin Instit.* 357, 4222–4241. doi: 10.1007/s10846-019-01079-x

Li, Y., and Chao, X. (2022). Distance-entropy: an effective indicator for selecting informative data. *Front. Plant Sci.* 12:818895. doi: 10.3389/fpls.2021.818895

Li, Z., and Jia, Y. (2009). Algebraic criteria for consensus problems of continuous-time networked systems. *Int. J. Control* 82, 643–658. doi: 10.1080/00207170802123895

Lian, Z., and Deshmukh, A. (2006). Performance prediction of an unmanned airborn vehicle multi-agent system. *Eur. J. Oper. Res.* 172, 680–695. doi: 10.1016/j.ejor.2004.10.015

Liang, T., Duan, B., Luo, X., Ma, Y., Yuan, Z., Zhu, R., et al. (2021). Identification of high nitrogen use efficiency phenotype in rice (*Oryza sativa* L.) through entire growth duration by unmanned aerial vehicle multispectral imagery. *Front. Plant Sci.* 12:740414. doi: 10.3389/fpls.2021.740414

Luo, X., Liao, J., Zang, Y., Zhou, Z., and Zang, Y. (2016). Improving agricultural mechanization level to promote agricultural sustainable development. *Trans. Chinese Soc. Agric. Eng.* 32, 1–11. doi: 10.11975/j.issn.1002-6819.2016.01.001

Ma, Z., Hu, F., and Yu, Z. (2006). "Multi-agent systems formal model for unmanned ground vehicles," in *2006 International Conference on Computational Intelligence and Security*, *Vol. 1* (Guangzhou), 492–497. doi: 10.1109/ICCIAS.2006.294183

Ogata, K. (1995). *Discrete-Time Control Systems*. New Jersey: Prentice-Hall, Inc.

Olfati-Saber, R., and Murray, R. M. (2004). Consensus problems in networks of agents with switching topology and time-delays. *IEEE Trans. Autom. Control* 49, 1520–1533. doi: 10.1109/TAC.2004.834113

Park, M. J., Kwon, O., and Seuret, A. (2016). Weighted consensus protocols design based on network centrality for multi-agent systems with sampled-data. *IEEE Trans. Autom. Control* 62, 2916–2922. doi: 10.1109/TAC.2016.2604682

Ran, G., Liu, J., Li, C., Chen, L., and Li, D. (2020). Event-based finite-time consensus control of second-order delayed multi-agent systems. *IEEE Trans. Circuits Syst. II* 68, 276–280. doi: 10.1109/TCSII.2020.2999480

Ren, W., and Beard, R. W. (2005). Consensus seeking in multiagent systems under dynamically changing interaction topologies. *IEEE Trans. Autom. Control* 50, 655–661. doi: 10.1109/TAC.2005.846556

Seuret, A., Liu, K., and Gouaisbaut, F. (2018). Generalized reciprocally convex combination lemmas and its application to time-delay systems. *Automatica* 95, 488–493. doi: 10.1016/j.automatica.2018.06.017

Sun, Y., and Wang, L. (2009). Consensus of multi-agent systems in directed networks with nonuniform time-varying delays. *IEEE Trans. Autom. Control* 54, 1607–1613. doi: 10.1109/TAC.2009.2017963

Teng, Y., Zhang, J., Dong, S., Zheng, S., and Liu, L. (2022). MSR-RCNN: a multi-class crop pest detection network based on a multi-scale super-resolution feature enhancement module. *Front. Plant Sci.* 13:810546. doi: 10.3389/fpls.2022.810546

Yan, Y., and Huang, J. (2017). Cooperative output regulation of discrete-time linear time-delay multi-agent systems under switching network. *Neurocomputing* 241, 108–114. doi: 10.1016/j.neucom.2017.02.022

Yang, M., Chen, C., and Kuo, Y. (2013). Implementation of intelligent air conditioner for fine agriculture. *Energy Build.* 60, 364–371. doi: 10.1016/j.enbuild.2013.01.034

Yu, C., Dong, Y., Li, Y., and Chen, Y. (2020). Distributed multi-agent deep reinforcement learning for cooperative multi-robot pursuit. *J. Eng.* 2020, 499–504. doi: 10.1049/joe.2019.1200

Zhang, D., Xu, H., Wang,H., Chen, X. (2022). Performance analysis for interconnected time-delay systems with networked communication. *ISA Trans.* doi: 10.1016/j.isatra.2022.02.012

Zou, W., Shi, P., Xiang, Z., and Shi, Y. (2019). Finite-time consensus of second-order switched nonlinear multi-agent systems. *IEEE Trans. Neural Netw. Learn. Syst.* 31, 1757–1762. doi: 10.1109/TNNLS.2019.2920880

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

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

frontiers | Frontiers in Plant Science

ORIGINAL RESEARCH
published: 18 July 2022
doi: 10.3389/fpls.2022.841452

Soluble Solids Content Binary Classification of Miyagawa Satsuma in Chongming Island Based on Near Infrared Spectroscopy

Yuzhen Chen[1,2†], Wanxia Sun[1†], Songtao Jiu[1], Lei Wang[1], Bohan Deng[1], Zili Chen[1], Fei Jiang[3], Menghan Hu[1,2] and Caixi Zhang[1*]*

[1] School of Agriculture and Biology, Shanghai Jiao Tong University, Shanghai, China, [2] Shanghai Key Laboratory of Multidimensional Information Processing, School of Communication and Electronic Engineering, East China Normal University, Shanghai, China, [3] Shanghai Citrus Research Institute, Shanghai, China

OPEN ACCESS

Edited by:
Chuanlei Zhang,
Tianjin University of Science and Technology, China

Reviewed by:
Ke Han,
Harbin University of Commerce, China
Xing Wei,
Purdue University, United States

***Correspondence:**
Menghan Hu
mhhu@ce.ecnu.edu.cn
Caixi Zhang
acaizh@sjtu.edu.cn

[†] These authors have contributed equally to this work

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

Received: 22 December 2021
Accepted: 15 June 2022
Published: 18 July 2022

Citation:
Chen Y, Sun W, Jiu S, Wang L, Deng B, Chen Z, Jiang F, Hu M and Zhang C (2022) Soluble Solids Content Binary Classification of Miyagawa Satsuma in Chongming Island Based on Near Infrared Spectroscopy. Front. Plant Sci. 13:841452. doi: 10.3389/fpls.2022.841452

Citrus is one of the most important fruits in China. Miyagawa Satsuma, one kind of citrus, is a nutritious agricultural product with regional characteristics of Chongming Island. Near-infrared Spectroscopy (NIR) is a proper method for studying the quality of fruits, because it is low-cost, efficient, non-destructive, and repeatable. Therefore, the NIR technique is used to detect citrus's soluble solid content (SSC) in this study. After obtaining the original spectral data, the first 70% of them are divided into the training set and 30% into the test set. Then, the Random Frog algorithm is chosen to select characteristic wavelengths, which reduces the dimension of the data and the complexity of the model, and accordingly makes the generalization of the classification model better. After comparing the performance of various classifiers (AdaBoost, KNN, LS-SVM, and Bayes) under different characteristic wavelength numbers, the AdaBoost classifier outperforms using 275 characteristic wavelengths for modeling eventually. The accuracy, precision, recall, and F_1-score are 78.3%, 80.5%, 78.3%, and 0.780, respectively and the ROC (Receiver Operating Characteristic Curve, ROC curve) is close to the upper left corner, suggesting that the classification model is acceptable. The results demonstrate that it is feasible to use the NIR technique to estimate whether the citrus is sweet or not. Furthermore, it is beneficial for us to apply the obtained models for identifying the quality of citrus correctly. For fruit traders, the model helps them to determine the growth cycle of citrus more scientifically, improve the level of citrus cultivation and management and the final fruit quality, and thus increase the economic income of fruit traders.

Keywords: near infrared spectroscopy, AdaBoost, random frog, citrus soluble solids content, machine learning

1. INTRODUCTION

Citrus fruits are among the most commonly grown and consumed fruits all over the world and meanwhile one of the most important fruits in China since they are very nutritious and can supplement vitamins, promote digestion and increase appetite (Zou et al., 2016; Anticona et al., 2020). The total output of *Citrus reticulata Blanco* is 21.2 million tons in China, accounting for 67% of the total citrus output. *Citrus unshiu* is one of the three main varieties of citrus reticulata

Blanco in China (Nam et al., 2019; Cheng et al., 2020). This research uses Miyagawa Satsuma, a variety of *citrus unshiu*, from Chongming Island in Shanghai, as the research object. The citrus in Chongming Island not only grows in environmental conditions famous for fresh air, clean water, and rich soil but also ripens in cultivation technology of "green prevention and control and plastic film covering with grass and organic cultivation" (this cultivation concept originates from Shanghai Qianwei Citrus Co., Ltd.). As a result, it owns the advantages of both rich nutrition and the regional characteristics of Chongming Island.

Soluble solids content (SSC), is one of the most important internal quality attributes of most fruits. The SSC plays an important role in the fruit maturity process and partly influences the flavor of most fruits, thus determining the acceptance of rich nutrients and economic benefits in the fruit trade. The detection of citrus SSC is not only beneficial to customers but also significant for growers (Li et al., 2016b; Fan et al., 2019; Guo et al., 2020). Therefore, in recent decades, the demand to develop non-destructive and rapid evaluation methods for citrus SSC has become more extensive and urgent. Electronic nose technology (Zhang et al., 2008, 2016), computer vision (Xia et al., 2016; Bhargava and Bansal, 2021), and hyperspectral imaging technology (Li et al., 2016a, 2018) are some common methods to measure the quality of fruits. However, electronic nose technology is restricted to limited enclosed space, which is inconvenient to carry out. Computer vision technology lacks spectral information. As for hyperspectral imaging technology, the obtained hypercube contains a lot of redundant information which leads to a high computation cost.

Fortunately, with the advantage of low testing cost, high efficiency, good reproducibility of test results, and non-destructive testing, NIR spectroscopy, between wavelength region range of 780–2,526 nm, has been applied popularly in the analysis of different fruit or vegetable samples (Beghi et al., 2017; Arendse et al., 2018), such as apple (Xia et al., 2020; Arefi et al., 2021; Ma et al., 2021; Li et al., 2022), tomato (Huang et al., 2021; Zhang et al., 2021), persimmon (Wei et al., 2020), pear (Cruz et al., 2021), and banana (Cruz et al., 2021). Xia et al. (2020) studied the effect of sample diameter differences on the online prediction of SSC of "Fuji" apples with the methods of visible and near-infrared spectroscopy and partial least square regression. It is justified that diffuse transmission spectra in 710–980 nm and diameter correction method with calculated attenuation coefficient are the best. Wei et al. used NIR hyperspectral imaging within 900–1,700 nm to model SSC and firmness determination of persimmon with partial least squares regression (Wei et al., 2020). The final models obtained a coefficient of determination of 0.757, RMSEP of 1.404 Brix, and R_p^2 of 0.876, RMSEP of 0.395 for SSC and firmness detection, respectively. Pahlawan et al. developed the calibration model to predict the SSC of bananas using NIR spectroscopy in the range from 350 to 1,000 nm. It was conducted by various distances of fiber optic probes to bananas samples (Cruz et al., 2021). To our best knowledge, there have been few similar studies on Miyagawa Satsuma. From these researches mentioned above, it can be easily seen that they focus on predicting the accurate number of the attribute focused on, such as SSC. Sometimes, we are more interested in knowing the sugar level rating rather than the specific value. There is no exact numerical index for distinguishing sweet and unsweetened, and therefore, we calculate the average value of SSC as the demarcation index for judging sweet or unsweet for the reason that SSC is an important index affecting the sweetness.

Reducing dimensions and seeking the most informative wavelengths are effective methods for processing data while selecting the most informative wavelengths of target information is an effective measure to simplify computation and improve the model performance (Li et al., 2019; Zhou et al., 2020). First, it has been shown that the inclusion of uninformative wavelengths while modeling affects the performance of predicting or classifying and model interpretability (Chang et al., 2016). Second, the identification of wavelengths that contain information about the attribute the research focuses on, will reduce the computation time and cost, from a more practical point of view (Zhang et al., 2019; Mamouei et al., 2020). Li et al. (2016a) chose the carlo-uninformative variable elimination and successive projections algorithm to select the most effective variables from hyperspectral data when doing the research on measuring SSC in pear. The results indicated that the model built using 18 effective variables achieved the optimal performance for the prediction of SSC. Jun et al. (2018) used an iteratively retaining informative variables algorithm to obtain 10 characteristic wavelengths when processing samples in predicting the SSC of cherry tomatoes. The experimental results showed the IRIV−CS−SVR model for SSC prediction could reach accuracy with R_p^2 = 0.9718 and R_c^2 = 0.9845. Fan et al. (2014) adopted a combination of the standard normal variate, uninformative variable elimination, genetic algorithm, and successive projections algorithm to obtain 30 characteristic wavelengths selected from full-spectra achieving the optimal performance.

In the current study, the binary classification of Miyagawa Satsuma is focused on, which owns the regional characteristics of Chongming Island. The classification model for nondestructive determination of Miyagawa Satsuma SSC will judge the quality of citrus more scientifically, and overcome the shortcomings of subjective differences and low efficiency. Meanwhile, it can identify the growth cycle of citrus and estimate the maturity time more accurately, which is conducive to the management arrangement such as picking. It can also provide a theoretical basis for citrus grading, which is good for fruit farmers or manufacturers to sell graded citrus and improve profits (Kundu et al., 2021).

2. MATERIALS AND METHODS

2.1. Data Collection

2.1.1. Near Infrared Spectra Acquisition

The equipment employed in our research is the Fourier transform near infrared spectrometer, an antaris II–F-NIR analyzer made by Thermo Fisher. We set NIR acquisition mode as integrating sphere mode and the gain as × 1. The NIR spectra are within the range from 1,000 nm to 2,500 nm.

All samples of Miyagawa Satsuma came from Shanghai Qianwei citrus Co., Ltd located on Chongming Island. All

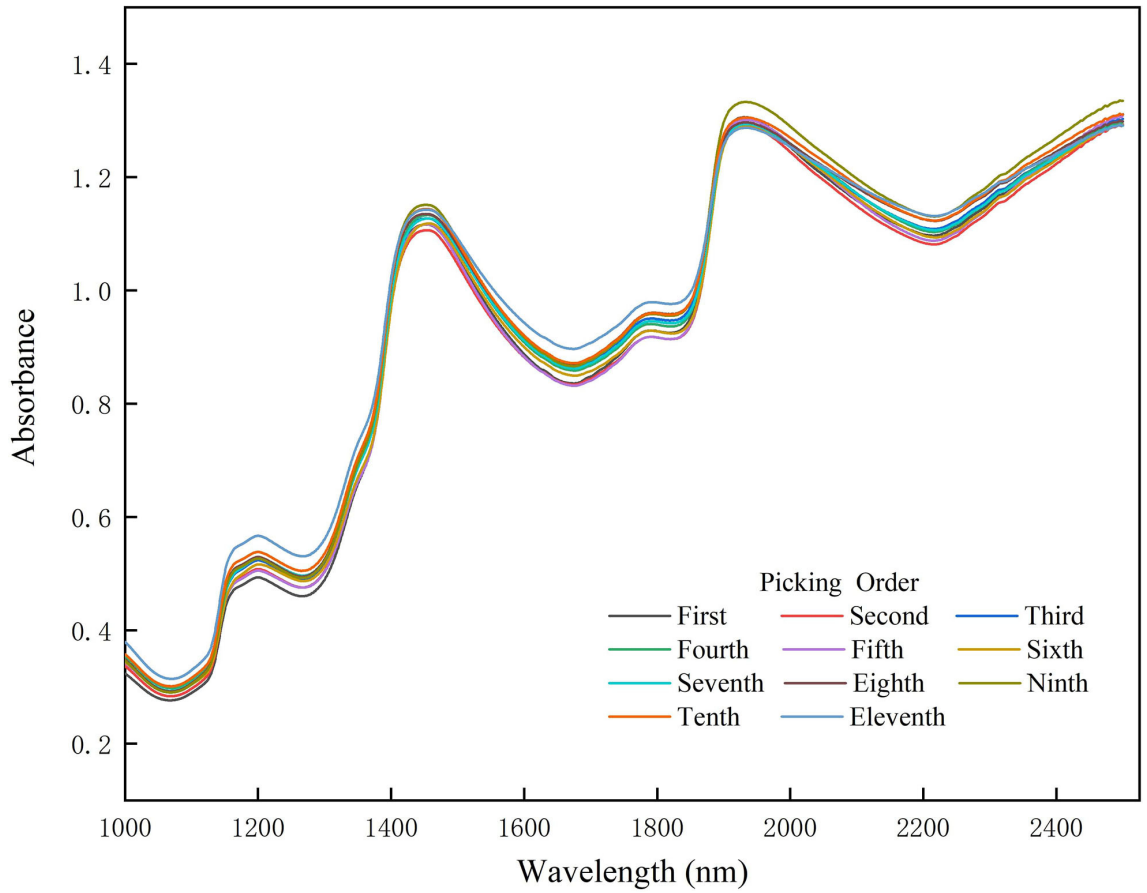


FIGURE 1 | Near infrared (NIR) spectra of citrus in different picking times in chronological order. Different picking orders are represented by different colors. Each line is the average spectrum of 12 fruit samples at each picking time.

samplings, 11 times total, were carried out within 3 months. In each sampling, five trees with the most similar growth were selected, which were without films and not among the outermost three rows of trees. Then the five trees were divided into upper, middle, and lower parts, where one sample was picked, respectively, from four directions: south, east, north, and west. As a result, 12 samples were obtained per tree, and a total of 60 were taken for each sampling. Next, 12 samples were randomly chosen among a total of 60 fruits. For each sample in the 12 fruits, the NIR spectra were gained from 4 points at the cross symmetry of the equatorial plane of the fruit. Finally, the averaged NIR spectra, obtained by averaging NIR spectra of four points, were taken as the original NIR spectra, as shown in **Figure 1**.

2.1.2. Soluble Solids Content Acquisition

After the Miyagawa Satsuma was squeezed and centrifuged, the SSC of the selected Miyagawa Satsuma samples was measured with a saccharometer (PR-101; Atago Co., Tokyo, Japan).

2.2. Data Preprocessing

The samples with obviously incomplete or wrong data are eliminated, whether NIR spectra or SSC, thus obtaining a total of 122 samples. Then samples were divided into the training set and test set by the SPXY algorithm (Galvao et al., 2005), with 70% of the samples as the training set and 30% as the test set. The principle of the Kennard stone algorithm (KS) algorithm is to calculate the Euclidean distance among all samples: select two samples with the maximum Euclidean distance into the training set, then carry out the iterative calculation, select the samples with the maximum and minimum Euclidean distance into the training set until the number of samples required by the training set is reached. SPXY algorithm is based on the KS algorithm, and it furthermore involves the chemical values and spectra among samples when calculating Euclidean distance, which makes the training set more representative, and makes the generalization ability of the established prediction model better.

2.3. Characteristic Wavelength Selection

The random frog (RF) algorithm (Li et al., 2012) was used to obtain the corresponding number of characteristic wavelengths of NIR spectral data, which has the features of conceptually simplicity, and fewer parameters to be trained in algorithm implementation, strong global search and optimization ability, etc. The principles of the algorithm are as follows. Each sample

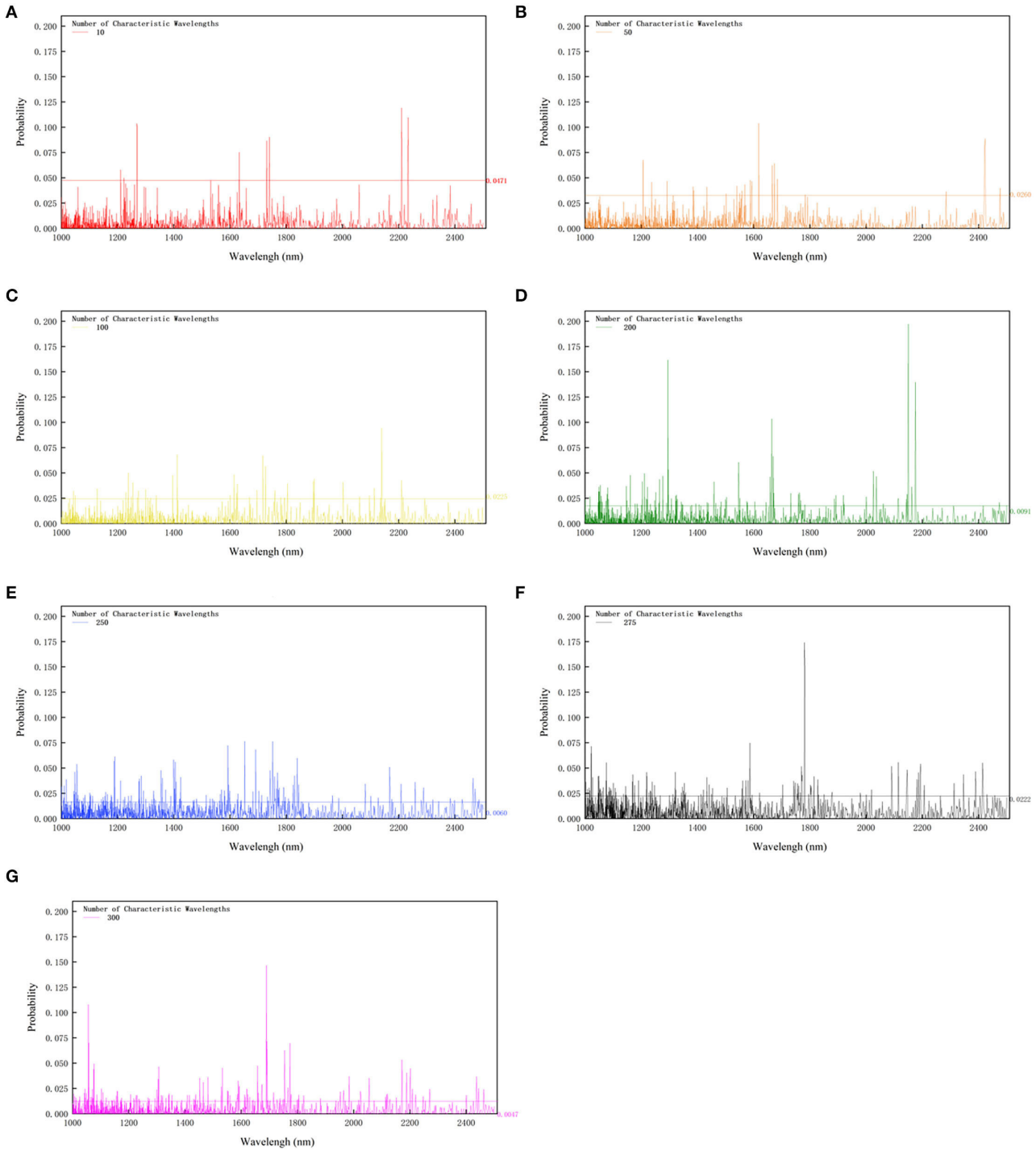


FIGURE 2 | Cutoff probability of different characteristic wavelengths. For the subsequent modeling, 10 **(A)**, 50 **(B)**, 100 **(C)**, 200 **(D)**, 250 **(E)**, 275 **(F)**, and 300 **(G)** characteristic wavelengths are chosen, respectively. The numbers on the right corner of the figure are the cutoff probability of corresponding characteristic wavelengths number. The cutoff probability generally decreases as the number of informative wavelengths increases.

in a population is regarded as a frog. Then the whole population is divided into m sub-groups with the scale of n. In each sub-group, the frogs with the best and worst fitness are used to produce a new child frog, which can be viewed as a jump of the best frog. If the fitness of the child frog is better than the parent with the worst fitness, replace the worst parent with

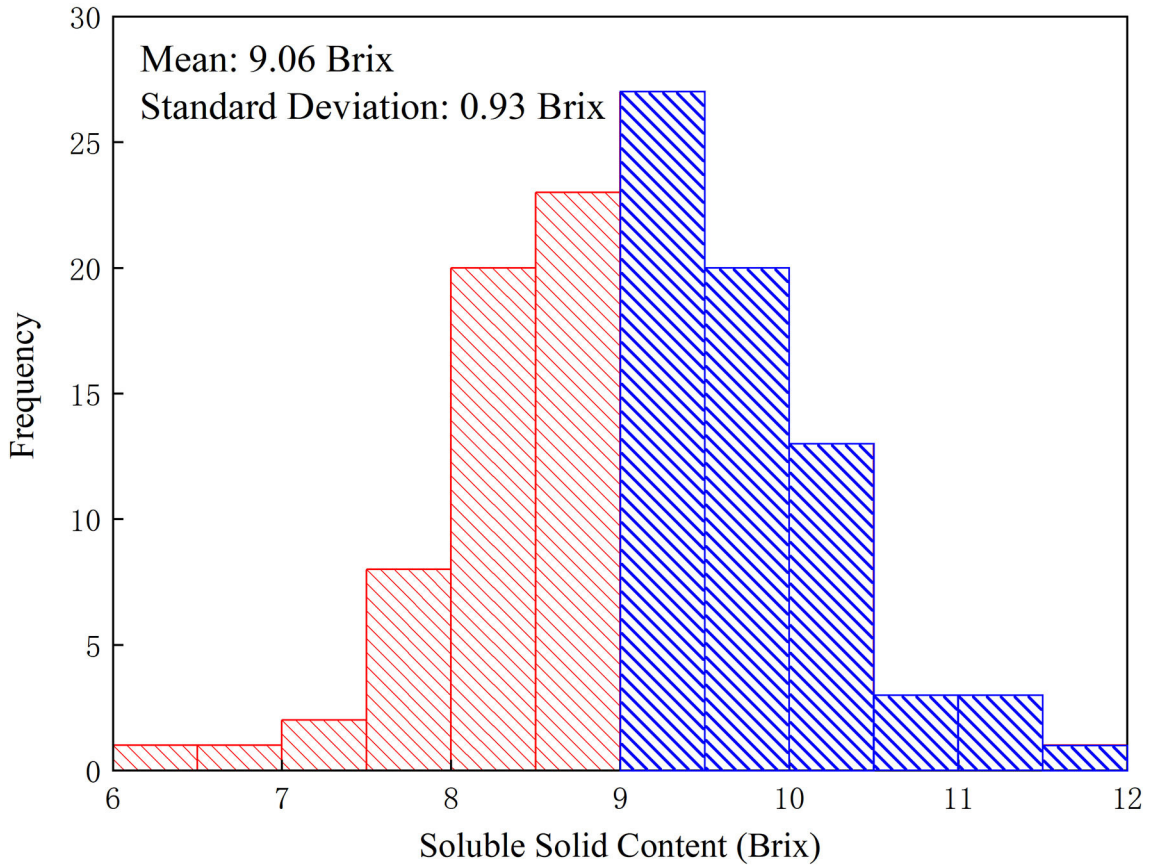


FIGURE 3 | Histogram of citrus soluble solid content (SSC) of the total 122 fruit samples collected in this study. The mean of all citruses' SSC is 9.06 Brix while the SD is 0.93 Brix. The citruses corresponding to the blue areas are regarded as sweet while the citruses corresponding to the red ones are regarded as unsweet.

the child, otherwise, randomly generate a new child, which can be viewed as the best frog's jumping again. If the fitness of the new child frog is still worse than the worst parent, then randomly generate another new child to replace the parent with the worst fitness. The evolutionary strategy of the random frog algorithm is like frogs jumping toward the optimal solution so that the algorithm gradually converges to the optimal solution.

The more specific steps of this algorithm are as follows: First, initialize parameters. Second, randomly generate an initial frog group and calculate the fitness of each frog. Third, arrange the frogs in descending order according to the value of fitness, and record the local optimal solution P_x. Then divide the F frogs from the initial group into sub-groups, namely, allocate F frogs into m sub-groups with the scale of n. Fourth, do a local search process, i.e., do the process described above in each sub-group. As a result, sub-groups do the fourth process, redivide the frog group, do the same operation as the first round, and record the global optimal solution P_x. Fifth, verify the calculation stop condition. If the convergence conditions of the algorithm are reached, the RF algorithm ends. If the global optimal solution has not been significantly improved, the execution of the algorithm should also be stopped.

To validate the performance of the RF algorithm in this task, the other common wavelength selection namely the competitive adaptive reweighted sampling algorithm (CARS) is used for comparison with the RF algorithm.

2.4. Binary Classification Model

The AdaBoost classifier (Freund et al., 1999) is selected for modeling. Boosting is an important integrated learning technology, which can enhance weak classifiers with poor prediction performance into strong classifications with good prediction performance in a cascade way. The core of its adaptability is that the wrong samples of the previous basic classifier will be strengthened, and all the weighted samples will be used to train the next basic classifier again. At the same time, a new weak classifier is added in each round until a predetermined small enough error rate or a predetermined maximum number of iterations is reached.

Specifically, the entire AdaBoost iterative algorithm consists of three steps: First, initialize the weight distribution of training data. If there are n samples, each training sample is given the same weight of $1/n$ at the beginning. Second, train weak classifiers. In the specific training process, if a sample point has been accurately classified, its weight will be reduced in the construction of the

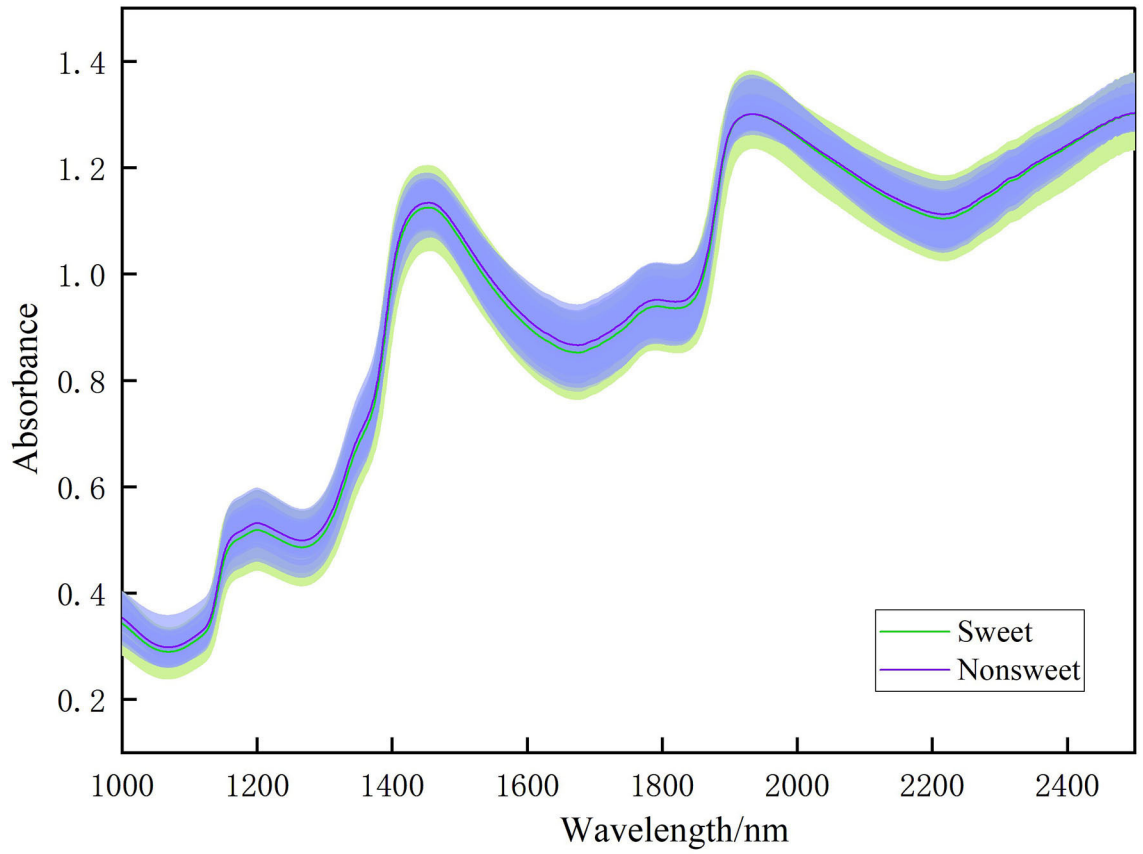


FIGURE 4 | Comparison of NIR spectra of sweet (SSC beyond 9 Brix) and unsweet (SSC below 9 Brix) fruit samples. The shaded areas represent the confident intervals to each line.

next training set; On the contrary, if a sample point is not accurately classified, its weight will be improved. Then, the weight updated sample set is used to train the next classifier, and the whole training process goes on in this way, iteratively. Third, combine the trained weak classifiers into strong classifiers. After the training process of each weak classifier, increase the weight of the weak classifier with a small classification error rate to make it play a greater decisive role in the final classification function while doing the opposite operation for the weak classifier with a large classification error rate.

To compare the performance of various classifiers, we choose AdaBoost, k-Nearest Neighbor (KNN), Bayes classifier, and LS-SVM to explore the best-performing classification model. In the current study, we use Matlab and Weka to establish the models.

2.5. Model Evaluation

To verify the efficiency of the classification system, evaluation indicators viz. confusion matrix, accuracy, precision, recall, F_1, micro-measures, and macro-measures are considered.

1) Confusion matrix: Assume that "Positive" means the positive samples and that "Negative" means the negative samples. Meanwhile, "True" represents that the prediction is right while "False" represents that the prediction is wrong. As a result, "TP" and "TN" mean that the positive sample is classified as "Positive" and that the negative sample is labeled as "Negative", respectively. "FP" and "FN" represent that the negative sample is labeled as "Positive" and that the positive sample is classified as "False." The four indicators make up the confusion matrix.

2) Accuracy: It is a ratio that is used to estimate the classification ability of a model within the range from 0 to 1. Generally speaking, the larger accuracy is, the better the classification is. It can be calculated by the following equation:

$$Accuracy = \frac{TN + TP}{TN + TP + FP + FN} \tag{1}$$

3) Precision: Precision is only used to evaluate the classification ability of the positive samples within the range from 0 to 1. It is obvious that the larger precision is, the more effective the system is. It is computed by:

$$Precision = \frac{TP}{TP + FP} \tag{2}$$

4) Recall: It is a ratio from 0 to 1. Obviously, the more it is close to 1, the better the system is. The calculation equation is:

$$Recall = \frac{TP}{TP + FN} \tag{3}$$

5) F_1: It is a harmonic mean of recall and precision. In this study, we consider the weight of recall and precision the same, which means attaching the weight of 0.5 to either of them. It is calculated by:

$$F_1 = \frac{2 * Precision * Recall}{Precision + Recall} \quad (4)$$

6) Receiver Operating Characteristic (ROC) Curve and Area Under Curve (AUC): The abscissa of the ROC curve is the false positive rate (FPR) while the ordinate is the true positive rate (TPR), where FPR $= \frac{FP}{TN+FP}$ and TPR $= \frac{YP}{TN+FN}$. Generally speaking, the closer the ROC curve is to the upper left corner of the image, the better the performance of the binary classifier. AUC is the area under the ROC curve and it is generally within the range of (0.5, 1). When the closer the ROC curve is to the upper left corner, the greater the value of AUC.

3. EXPERIMENTAL RESULTS AND ANALYSIS

3.1. NIR Spectral Characteristics of Miyagawa Satsuma in Different Picking Time

The NIR spectra in different picking times are shown in **Figure 1**. The trend of the citrus spectra collected each time is similar. There is an obvious absorption trough near 1,080, 1,300, 1,700, and 2,200 nm, respectively. According to the principles of NIR spectroscopy, due to the fact that the sample will selectively absorb NIR waves with different frequencies, the NIR wave which passes through the sample will become weaker in some wavelength ranges, and the transmitted NIR wave will carry the information of organic component and structure. Therefore, it can be inferred that these absorption troughs can probably be the most informative areas, which can be reflected in the characteristic wavelength selection.

Our reason for using fruits with different picking periods for modeling is to increase the coverage of the SSC, allowing a larger range of variation in the spectral data and ultimately increasing the model robustness. We performed the statistical tests on the obtained spectral data and SSC and found significant differences between spectral data and SSC for non-adjacent picking periods ($p < 0.05$) and no significant differences for adjacent picking periods ($p > 0.05$). This is in accordance with expectations. Because, as the fruit ripens, the SSC will certainly increase and the spectral differences will increase.

3.2. Performance of RF

As mentioned above, the characteristic wavelength selection can accelerate the computation speed and reduce computation cost to a degree. RF algorithm is chosen to generate characteristic wavelengths with the numbers 10, 50, 100, 200, 250, 275, and 300, respectively, which is displayed in **Figure 2**. It is easy to find that the larger the number of characteristic wavelengths is, the smaller the cutoff probability is. The cutoff probability indicates the threshold value for screening the required number of informative wavelengths. The wavelength numbers, 10, 50, 100, 200, 250, 275, and 300, respectively correspond to the cutoff probabilities, 0.0471, 0.0260, 0.0225, 0.0222, 0.0091, 0.0060, and 0.0047. The cutoff probability generally decreases as the number of informative wavelengths increases (**Figure 2**). Meanwhile, it is true with what has been inferred in the above section that the absorption troughs can be the most informative, most of the retaining wavelengths gather in the areas inferred before viz. 1,080, 1,300, 1,700, and 2,200 nm. This probably has a relationship with the functional groups viz. —OH, —CH, —NH.

In addition, the classification models based on CARS selected wavelengths are established, and their performance is not as good as the RF-based models. For example, when ten characteristic wavelengths are selected, the RF-based model gives a better

TABLE 1 | Modeling results of sweet (SSC beyond 9 Brix) and unsweet (SSC below 9 Brix) classification of Miyagawa Satsuma in Chongming Island under different classifiers with different characteristic wavelengths.

Characteristic wavelengths	Models	Metrics			
		Accuracy (%)	Precision (%)	Recall (%)	F1-score
10	AdaBoost	60.9	62.1	60.9	0.604
	KNN	69.6	69.8	69.6	0.696
	Bayes	65.2	65.4	65.2	0.648
	LS-SVM	62.2	53.3	100.0	0.696
50	AdaBoost	60.9	62.1	60.9	0.604
	KNN	52.2	53.7	52.2	0.503
	Bayes	65.2	65.4	65.2	0.648
	LS-SVM	67.6	57.1	100.0	0.727
100	AdaBoost	69.6	69.8	69.6	0.696
	KNN	52.2	52.9	52.2	0.516
	Bayes	69.6	69.6	69.6	0.694
	LS-SVM	59.5	51.6	100.0	0.681
200	AdaBoost	65.2	71.6	65.2	0.631
	KNN	65.2	67.8	65.2	0.644
	Bayes	69.6	69.6	69.6	0.694
	LS-SVM	62.2	53.3	100.0	0.696
250	AdaBoost	69.6	71.3	69.6	0.692
	KNN	56.5	58.1	56.5	0.555
	Bayes	69.6	69.8	0.7	0.696
	LS-SVM	62.2	53.3	100.0	0.696
275	**AdaBoost**	**78.3**	**80.5**	**78.3**	**0.780**
	KNN	60.9	62.1	60.9	0.604
	Bayes	69.6	69.8	69.6	0.696
	LS-SVM	62.2	53.3	100.0	0.696
300	AdaBoost	69.6	71.3	69.6	0.692
	KNN	65.2	67.8	65.2	0.644
	Bayes	69.6	69.8	69.6	0.696
	LS-SVM	62.2	54.2	81.3	0.65
1556	AdaBoost	75.0	68.8	91.7	0.786
	KNN	60.9	56.3	81.8	0.667
	Bayes	73.9	72.7	72.7	0.727
	LS-SVM	56.8	50.0	93.8	0.652

Bold font represents the best model.

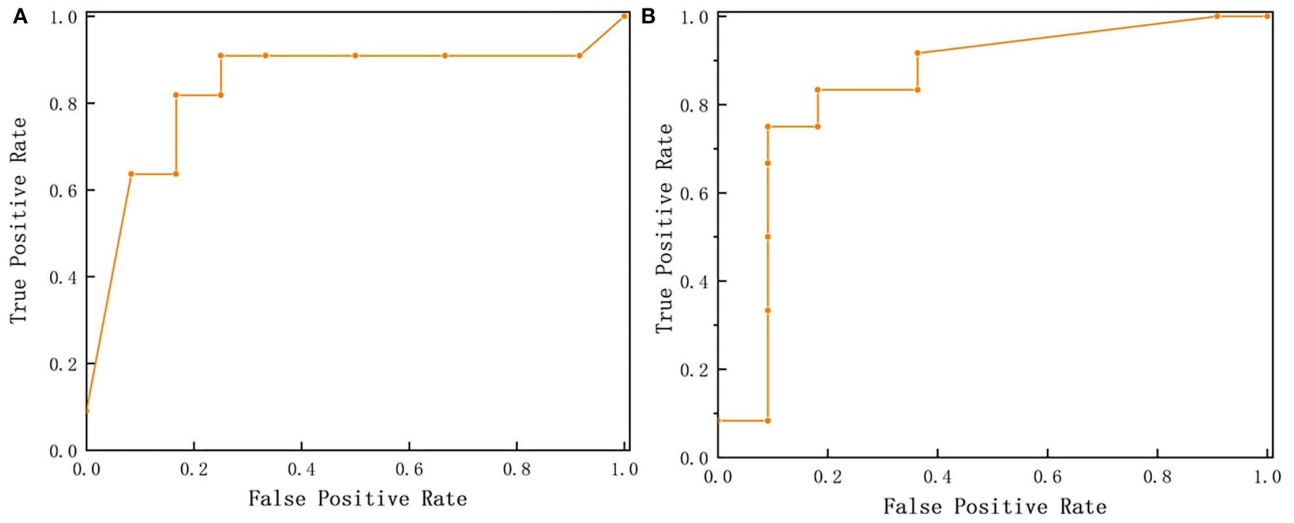


FIGURE 5 | Receiver Operating Characteristic (ROC) curves. **(A)** is ROC curve of positive samples (SSC beyond 9 Brix) while **(B)** is ROC curve of negative samples (SSC below 9 Brix).

performance than the model based on CARS, with the accuracy of 60.9, 69.6, 65.2, and 62.2% vs. 52.78, 52.78, 58.33, and 47.22% for AdaBoost, KNN, Bayes, and LS-SVM modeling methods, respectively. Overall, CARS does not perform as well as RF for the informative wavelength selection.

3.3. Model Analysis Using Plant Physiology Phenomenon

The spectral properties of plants are mainly determined by their internal structure. For the current study, the obtained spectra are the result of the interaction of the incident light with the chemical composition and physical structure of Citrus. For Miyagawa Satsuma, its structure can be divided into exocarp (oil cell layer), mesocarp (white cortex), endocarp, fruit, and fruit stem from outside to inside. Among them, the surface of the soluble dietary fiber of mandarin pulp is not smooth, the strips and gaps are intertwined, and there are raised particles; the surface of the soluble dietary fiber of mandarin peel is larger, but the surface depressions are mixed with a few spherical particles. There is a strong interaction between the two molecules.

As an important indicator for evaluating fruit sweetness, SSC is mainly composed of soluble sugars (including sucrose, fructose, and glucose). In the NIR region, the stretching and deformation vibration absorption peaks of $O-H$ bonds in soluble sugars are located around 1,440 and 2,080 nm, and there are three absorption peaks of soluble solids at 980, 1,169, and 1,485 nm (Musingarabwi et al., 2016). The water content has a great influence on the absorption of the plant spectrum. Under the condition of multi-layer leaves, the water absorption bands at 1,100 and 960 nm have a great influence on spectral reflectance. Absorption leads to a decrease in reflectance and an increase in absorbance, and peaks of reflectance (i.e., peaks and valleys of absorbance) appear at 1,600 and 2,200 nm (Ma et al., 2017).

The wavelengths selected by RF include three characteristic wavelengths near the absorption peaks of soluble solids at 980 nm, 1,169 nm, and 1,485 nm, and two characteristic wavelengths near the absorption peaks of stretching and deformation vibrations of $O-H$ bonds in soluble sugars at 1,440 nm and 2,080 nm, and a characteristic wavelength near the strong absorption peak of water at 1,400 nm. This analysis explains why the model based on the RF selected wavelengths performs better.

3.4. Soluble Solids Content Division

The research holds the opinion that consumers are more concerned about whether the Miyagawa Satsuma is sweet or not, but not the concrete value of sweetness. Referring to **Figure 3**, the dichotomous map or histogram of 122 Miyagawa Satsuma citruses' SSC, the distribution of this figure is roughly similar to the normal distribution, whose mean of all citruses' SSC is 9.06 Brix and the SD is 0.93 Brix. To carry out our belief, 9 Brix was taken as the boundary after asking an expert in agriculture for advice. As a consequence, the citruses with SSC more than or equal to 9 Brix are considered to be sweet and the others are not sweet for the following classification modeling.

Figure 4 shows the comparison of NIR spectra of sweet and unsweet fruit samples. As shown in **Figure 4**, the large overlap between the spectral curves of the sweet and unsweet samples indicates that the model will not perform as expected if the model is constructed based on original spectra. Therefore, we need to select the informative wavelengths specific to SSC classification, and then combine them with pattern recognition methods for modeling.

3.5. Performance Analysis of Different Binary Classification Models

As mentioned before, AdaBoost, KNN, Bayes, and LS-SVM are adopted to establish classification models. The performance comparison of different classifiers under different characteristic wavelengths is shown in **Table 1**. The conclusion can be drawn that when the number of characteristic wavelengths is 275, the classification model established by the AdaBoost classifier

performs best (bold in the table), with accuracy, precision, recall, and F_1-score 78.3%, 80.5%, 78.3%, 0.780, respectively.

From the perspective of the number of characteristic wavelengths, when the number is 10, the best performer is the KNN classifier, with accuracy, precision, recall, and F_1-score 69.6%, 69.8%, 69.6%, and 0.696. When the number is 50, LS-SVM performs best according to the accuracy of 67.6%, precision of 57.1%, recall of 100%, and F_1-score 0.727. When the number is 100, Adaboost performs best while the best performer belongs to Bayes when the number is 200. As for the number 250, the results of AdaBoost are as good as Bayes. Finally, AdaBoost still stands out among four classifiers when under the condition of 300 characteristic wavelengths. From **Table 1**, it can be found that when the number of characteristic wavelengths is either too small or too large, the performance of every different classifier is not as good as the situation when the number is proper, from the perspective of four classifiers.

Compared to the results of original wavelengths number 1,556 without any procession, the best results of AdaBoost, KNN, and LS-SVM happen when they are through characteristic wavelengths selection, however, except Bayes. But after weighing the wavelength reduction and performance, it is reasonable to think that characteristic wavelength selection also works for Bayes.

The ROC of positive and negative samples of the test set is shown in **Figure 5**. It can be seen that the ROC curves of positive and negative samples are all close to the upper left corner, and the total AUC is 0.841, indicating that the model has good robustness and can adapt flexibly to the uneven distribution of positive and negative samples in actual situations.

Too many spectral features bring information redundancy, and too few spectral features bring information loss. Based on the experimental results, for this classification task, the optimal number of spectral features is 275. Compared to the other modeling methods, the AdaBoost method achieves the best performance at 275 wavelength numbers. This is because AdaBoost combines multiple weak classifiers in a reasonable way to make one strong classifier. The other three methods used in this paper just give one separate model.

4. CONCLUSION AND REFLECTION

Based on NIR spectroscopy, the random frog algorithm, and AdaBoost algorithm, and taking citrus in Shanghai Chongming Island as the research object, this study focuses on the problems of binary classification between NIR spectra and Miyagawa Satsuma SSC. Nine Brix is selected as the threshold of being sweet or not and the samples are divided into the training set and test set. After selecting characteristic wavelengths through the RF algorithm, they are used to establish binary classification models by AdaBoost, LS-SVM, and other classifiers. According to their performance, the AdaBoost classifier is the optimum model, with accuracy, precision, recall, and F_1-score 78.3%, 80.5%, 78.3%, and 0.780, respectively.

Analyzing the model performance, we find that the constructed model does not have a very high performance. Combined with the sampling process and the test results, two reasons may be summarized (1) due to the limited penetration depth of NIR and the thick skin of the fruit, most of the NIR light does not penetrate the skin to reach the fruit part; and (2) there are environment disturbances during sampling and instrument errors in the process of collecting spectra.

The constructed model has the potential to be embedded in portable NIR acquisition devices in the future, which can facilitate fruit farmers to judge the quality of the citrus and be conducive to improving the sale pricing system of citrus in Chongming Island, so as to maximize the sale profit of fruit-sellers.

DATA AVAILABILITY STATEMENT

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

AUTHOR CONTRIBUTIONS

CZ and MH: funding acquisition. YC, MH, WS, SJ, LW, BD, ZC, and CZ: methodology and validation. YC, WS, and MH: writing—original draft. YC, SJ, LW, MH, and CZ: writing—review and editing. FJ: providing citrus materials. All authors contributed to the article and approved the submitted version.

FUNDING

This study is sponsored by Shanghai Agriculture Applied Technology Development Program (Grant No. X20200102), Shanghai Agricultural System Standard Development Program (2018-013), Agriculture Research System of Shanghai (Grant No. 201407), Shanghai Agriculture Applied Technology Development Program (Grant No. T20220103), the Science and Technology Commission of Shanghai Municipality (No. 19511120100), the GHfund B (No. 20210702), and the Fundamental Research Funds for the Central Universities.

REFERENCES

Anticona, M., Blesa, J., Frigola, A., and Esteve, M. J. (2020). High biological value compounds extraction from citrus waste with non-conventional methods. *Foods* 9, 811. doi: 10.3390/foods9060811

Arefi, A., Sturm, B., von Gersdorff, G., Nasirahmadi, A., and Hensel, O. (2021). Vis-nir hyperspectral imaging along with gaussian process regression to monitor quality attributes of apple slices during drying. *LWT* 152, 112297. doi: 10.1016/j.lwt.2021.112297

Arendse, E., Fawole, O. A., Magwaza, L. S., and Opara, U. L. (2018). Non-destructive prediction of internal and external quality attributes of fruit with thick rind: a review. *J. Food Eng.* 217, 11–23. doi: 10.1016/j.jfoodeng.2017.08.009

Beghi, R., Buratti, S., Giovenzana, V., Benedetti, S., and Guidetti, R. (2017). Electronic nose and visible-near infrared spectroscopy in fruit and

vegetable monitoring. *Rev. Anal. Chem*. 36, 1–24. doi: 10.1515/revac-2016-0016

Bhargava, A., and Bansal, A. (2021). Fruits and vegetables quality evaluation using computer vision: a review. *J. King Saud Univer. Comput. Inf. Sci*. 33, 243–257. doi: 10.1016/j.jksuci.2018.06.002

Chang, H., Zhu, L., Lou, X., Meng, X., Guo, Y., and Wang, Z. (2016). Local strategy combined with a wavelength selection method for multivariate calibration. *Sensors* 16, 827. doi: 10.3390/s16060827

Cheng, C.-,x., Jia, M., Gui, Y., and Ma, Y. (2020). Comparison of the effects of novel processing technologies and conventional thermal pasteurisation on the nutritional quality and aroma of mandarin (citrus unshiu) juice. *Innovat. Food Sci. Emerg. Technol*. 64, 102425. doi: 10.1016/j.ifset.2020.102425

Cruz, S., Guerra, R., Brazio, A., Cavaco, A. M., Antunes, D., and Passos, D. (2021). Nondestructive simultaneous prediction of internal browning disorder and quality attributes in 'rocha'pear (*Pyrus communis* L.) using vis-nir spectroscopy. *Postharvest Biol. Technol*. 179, 111562. doi: 10.1016/j.postharvbio.2021.111562

Fan, S., Huang, W., Li, J., Zhao, C., and Zhang, B. (2014). Characteristic wavelengths selection of soluble solids content of pear based on nir spectral and ls-svm. *Guang pu xue yu guang pu fen xi= Guang pu* 34, 2089–2093. doi: 10.3964/j.issn.1000-0593(2014)08-2089-05

Fan, S., Li, J., Xia, Y., Tian, X., Guo, Z., and Huang, W. (2019). Long-term evaluation of soluble solids content of apples with biological variability by using near-infrared spectroscopy and calibration transfer method. *Postharvest Biol. Technol*. 151, 79–87. doi: 10.1016/j.postharvbio.2019.02.001

Freund, Y., Schapire, R., and Abe, N. (1999). A short introduction to boosting. *J. Jpn. Soc. Artif. Intell*. 14, 1612.

Galvao, R. K. H., Araujo, M. C. U., José, G. E., Pontes, M. J. C., Silva, E. C., and Saldanha, T. C. B. (2005). A method for calibration and validation subset partitioning. *Talanta* 67, 736–740. doi: 10.1016/j.talanta.2005.03.025

Guo, Z., Wang, M., Agyekum, A. A., Wu, J., Chen, Q., Zuo, M., et al. (2020). Quantitative detection of apple watercore and soluble solids content by near infrared transmittance spectroscopy. *J. Food Eng*. 279, 109955. doi: 10.1016/j.jfoodeng.2020.109955

Huang, Y., Dong, W., Chen, Y., Wang, X., Luo, W., Zhan, B., et al. (2021). Online detection of soluble solids content and maturity of tomatoes using vis/nir full transmittance spectra. *Chemometr. Intell. Lab. Syst*. 210, 104243. doi: 10.1016/j.chemolab.2021.104243

Jun, S., Yating, L., Xiaohong, W., Chunxia, D., and Yong, C. (2018). Ssc prediction of cherry tomatoes based on iriv-cs-svr model and near infrared reflectance spectroscopy. *J. Food Process. Eng*. 41, e12884. doi: 10.1111/jfpe.12884

Kundu, N., Rani, G., Dhaka, V. S., Gupta, K., Nayak, S. C., Verma, S., et al. (2021). Iot and interpretable machine learning based framework for disease prediction in pearl millet. *Sensors* 21, 5386. doi: 10.3390/s21165386

Li, H.-D., Xu, Q.-S., and Liang, Y.-Z. (2012). Random frog: an efficient reversible jump markov chain monte carlo-like approach for variable selection with applications to gene selection and disease classification. *Anal. Chim. Acta* 740, 20–26. doi: 10.1016/j.aca.2012.06.031

Li, J., Tian, X., Huang, W., Zhang, B., and Fan, S. (2016a). Application of long-wave near infrared hyperspectral imaging for measurement of soluble solid content (ssc) in pear. *Food Anal. Methods* 9, 3087–3098. doi: 10.1007/s12161-016-0498-2

Li, J., Wang, Q., Xu, L., Tian, X., Xia, Y., and Fan, S. (2019). Comparison and optimization of models for determination of sugar content in pear by portable vis-nir spectroscopy coupled with wavelength selection algorithm. *Food Anal. Methods* 12, 12–22. doi: 10.1007/s12161-018-1326-7

Li, J.-L., Sun, D.-W., and Cheng, J.-H. (2016b). Recent advances in nondestructive analytical techniques for determining the total soluble solids in fruits: a review. *Comprehensive Rev. Food Sci. Food Safety* 15, 897–911. doi: 10.1111/1541-4337.12217

Li, L., Huang, W., Wang, Z., Liu, S., He, X., and Fan, S. (2022). Calibration transfer between developed portable vis/nir devices for detection of soluble solids contents in apple. *Postharvest Biol. Technol*. 183, 111720. doi: 10.1016/j.postharvbio.2021.111720

Li, X., Wei, Y., Xu, J., Feng, X., Wu, F., Zhou, R., et al. (2018). Ssc and ph for sweet assessment and maturity classification of harvested cherry fruit based on nir hyperspectral imaging technology. *Postharvest Biol. Technol*. 143, 112–118. doi: 10.1016/j.postharvbio.2018.05.003

Ma, T., Li, X., Inagaki, T., Yang, H., and Tsuchikawa, S. (2017). Noncontact evaluation of soluble solids content in apples by near-infrared hyperspectral imaging. *J. Food Eng*. 224, 53–61. doi: 10.1016/j.jfoodeng.2017.12.028

Ma, T., Xia, Y., Inagaki, T., and Tsuchikawa, S. (2021). Rapid and nondestructive evaluation of soluble solids content (ssc) and firmness in apple using vis-nir spatially resolved spectroscopy. *Postharvest Biol. Technol*. 173, 111417. doi: 10.1016/j.postharvbio.2020.111417

Mamouei, M., Budidha, K., Baishya, N., Qassem, M., and Kyriacou, P. (2020). Comparison of wavelength selection methods for in-vitro estimation of lactate: a new unconstrained, genetic algorithm-based wavelength selection. *Sci. Rep*. 10, 1–12. doi: 10.1038/s41598-020-73406-4

Musingarabwi, D. M., Nieuwoudt, H. H., Young, P. R., Eyéghè-Bickong, H., and Vivier, M. A. (2016). A rapid qualitative and quantitative evaluation of grape berries at various stages of development using fourier-transform infrared spectroscopy and multivariate data analysis. *Food Chem*. 190, 253–262. doi: 10.1016/j.foodchem.2015.05.080

Nam, H.-A., Ramakrishnan, S. R., and Kwon, J.-H. (2019). Effects of electron-beam irradiation on the quality characteristics of mandarin oranges (citrus unshiu (swingle) marcov) during storage. *Food Chem*. 286, 338–345. doi: 10.1016/j.foodchem.2019.02.009

Wei, X., He, J., Zheng, S., and Ye, D. (2020). Modeling for ssc and firmness detection of persimmon based on nir hyperspectral imaging by sample partitioning and variables selection. *Infrared Phys. Technol*. 105, 103099. doi: 10.1016/j.infrared.2019.103099

Xia, Y., Fan, S., Li, J., Tian, X., Huang, W., and Chen, L. (2020). Optimization and comparison of models for prediction of soluble solids content in apple by online vis/nir transmission coupled with diameter correction method. *Chemometr. Intell. Labo. Syst*. 201, 104017. doi: 10.1016/j.chemolab.2020.104017

Xia, Z., Wu, D., Nie, P., and He, Y. (2016). Non-invasive measurement of soluble solid content and ph in kyoho grapes using a computer vision technique. *Anal. Methods* 8, 3242–3248. doi: 10.1039/C5AY02694F

Zhang, D., Xu, Y., Huang, W., Tian, X., Xia, Y., Xu, L., et al. (2019). Nondestructive measurement of soluble solids content in apple using near infrared hyperspectral imaging coupled with wavelength selection algorithm. *Infrared Phys. Technol*. 98, 297–304. doi: 10.1016/j.infrared.2019.03.026

Zhang, D., Yang, Y., Chen, G., Tian, X., Wang, Z., Fan, S., et al. (2021). Nondestructive evaluation of soluble solids content in tomato with different stage by using vis/nir technology and multivariate algorithms. *Spectrochim. Acta A* 248, 119139. doi: 10.1016/j.saa.2020.119139

Zhang, H., Wang, J., and Ye, S. (2008). Predictions of acidity, soluble solids and firmness of pear using electronic nose technique. *J. Food Eng*. 86, 370–378. doi: 10.1016/j.jfoodeng.2007.08.026

Zhang, W., Pan, L., Zhao, X., and Tu, K. (2016). A study on soluble solids content assessment using electronic nose: persimmon fruit picked on different dates. *Int. J. Food Propert*. 19, 53–62. doi: 10.1080/10942912.2014.940535

Zhou, Q., Huang, W., Fan, S., Zhao, F., Liang, D., and Tian, X. (2020). Non-destructive discrimination of the variety of sweet maize seeds based on hyperspectral image coupled with wavelength selection algorithm. *Infrared Phys. Technol*. 109, 103418. doi: 10.1016/j.infrared.2020.103418

Zou, Z., Xi, W., Hu, Y., Nie, C., and Zhou, Z. (2016). Antioxidant activity of citrus fruits. *Food Chem*. 196, 885–896. doi: 10.1016/j.foodchem.2015.09.072

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

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

frontiers | Frontiers in Plant Science

TYPE Original Research
PUBLISHED 22 July 2022
DOI 10.3389/fpls.2022.897475

OPEN ACCESS

EDITED BY
Chuanlei Zhang,
Tianjin University of Science and Technology, China

REVIEWED BY
Ran Wang,
Beijing Academy of Agricultural and Forestry Sciences, China
Klaus H. Hoffmann,
University of Bayreuth, Germany
Varvara Yu. Vedenina,
Institute for Information Transmission Problems (RAS), Russia

*CORRESPONDENCE
Qing Yao
q-yao@126.com

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

RECEIVED 16 March 2022
ACCEPTED 05 July 2022
PUBLISHED 22 July 2022

CITATION
Feng Z, Wei Q, Ye Z, Yang B, Gao Y, Lv J, Dai Y, Bao J and Yao Q (2022) Vibrational courtship disruption of *Nilaparvata lugens* using artificial disruptive signals. *Front. Plant Sci.* 13:897475. doi: 10.3389/fpls.2022.897475

Vibrational courtship disruption of *Nilaparvata lugens* using artificial disruptive signals

Zelin Feng[1], Qi Wei[2], Zhongru Ye[1], Baojun Yang[2], Yufan Gao[1], Jun Lv[1], Yanyun Dai[1], Jia Bao[1] and Qing Yao[1]*

[1]School of Information Science and Technology, Zhejiang Sci-Tech University, Hangzhou, China, [2]State Key Laboratory of Rice Biology, China National Rice Research Institute, Hangzhou, China

The brown planthopper (BPH), *Nilaparvata lugens* (Stål; Hemiptera: Delphacidae) is a piercing-sucking pest that causes serious damage to rice plants by sucking the phloem sap from the plants and transmitting viruses. During courtship, the BPH vibrates its abdomen to produce signals that are transmitted to rice plants through its legs. Male BPHs search, locate, and mate with female BPHs after they exchange courtship signals with each other. Currently, spraying chemical pesticides is still the primary method for controlling BPH populations in paddy fields, although this approach has led to severe environmental pollution. A physical control method based on BPH courtship disruption to reduce the mating rate is a promising strategy for cutting environmental pollution. To acquire effective courtship disruptive signals, we developed a vibration signal recording, monitoring, and playback system for BPHs. Using this system, BPH courtship signals and male competition signals were collected and analyzed to obtain their frequency spectra. Results show that the mean main vibration frequency of female courtship signals is 234Hz and the mean pulse rate is 23Hz. The mean main vibration and pulse frequencies of the male courtship signals are 255Hz and 82Hz, respectively. Besides, the mean main vibration frequency of the male competition signal is 281Hz. Seven different forms and frequencies of artificial signals were played back to male BPHs, then the courtship and behavioral responses of male BPHs were analyzed. Results indicate that a pure tone of 225Hz prevents the males from recognizing female courtship signals. The male reply rate fell from 95.6 to 33.3% and the mean reply delay time increased from 5.3s to 9.1s. The reply rates of the other six artificial signals ranged from 42.9 to 83.7%, and the mean reply delays were between 5.0s and 9.3s. Therefore, the courtship behavior of BPHs can be disrupted by using specific artificial disruptive signals.

KEYWORDS

***Nilaparvata lugens*, courtship signals, disruptive signals, courtship disruption, behavior response**

Introduction

The brown planthopper (BPH), *Nilaparvata lugens* (Stål; Hemiptera: Delphacidae), is an agricultural pest that feeds on rice sap using piercing-sucking mouthparts (Figure 1). BPH outbreaks lead to widespread plant death, which directly affects rice yield and quality (Dyck and Thomas, 1979). At present, BPH control mainly relies on chemical pesticides (Wu, 2018), but the extensive use of pesticides often leads to several issues including pesticide residue, death of natural enemies, pest resistance, and environmental pollution (Matsumura and Sanada-Morimura, 2010). Therefore, the exploration of new, efficient, and environmentally-friendly BPH physical control technologies to replace chemical pesticides has become the focus of much research in recent years.

Ossiannilsson (1949) demonstrated that most groups of Auchenorrhyncha generate low-intensity vibration signals and use them for communication. Ichikawa et al. (1975) recorded the courtship signals of three species of rice planthoppers, including BPH. The BPH vibration signal cannot be heard by the human ear and can only be gathered using specialized devices (Zhang and Chen, 1987b). Zhang and Chen (1991) developed a BPH vibration signal recording device for BPH vibration signals with a pickup, preamplifier, power amplifier, and recorder. The identified vibration signals of BPHs mainly occur during courtship and mating (Zhang and Chen, 1987a; Saxena et al., 1993), and the vibration signals of BPHs have three roles. These roles include the communication and identification of male and female individuals, stimulation of sexual arousal, and guidance for males in search of females (Zhang, 1997). On the same or neighboring rice plants, BPHs use vibrating courtship signals to contact the opposite sex.

FIGURE 1
Nilaparvata lugens (adult).

Females generally stay in one place, while males call and search according to the response signals (Zhang, 1997). At the same time, BPH males are capable of emitting vibration signals related to reproductive competition, which disturb the other males and reduce their mating rate (Zhang et al., 1991). Since the courtship vibration signals of BPHs are highly specific and stable (Cokl and Virant-Doberlet, 2003), new pest control technologies can be utilized to emit disruptive signals that upset the courtship communication of the target pest (Eriksson et al., 2012; Polajnar et al., 2016).

Mazzoni et al. (2009) used artificially simulated competition signals to disrupt the male and female mating behavior of *Scaphoideus titanus* and proposed a new vibration control technology for leafhoppers. This technology used a shaker device to generate the selected disruptive signal and transmit the signal to grape leaves through trellising wires. The amplitude of the signal was 7.5 mm at the source and dropped to 0.1 mm at a distance of 10 m from the sound source. The mating inhibition rate of the male/female leafhoppers was about 90% and there was no negative impact on the natural enemies of the leafhoppers in the vineyard (Polajnar et al., 2016). Fu et al. (1997, 1999) confirmed that the competition signal of males significantly reduced the mating rate of BPHs (41.0%) and had a significant inhibitory effect during reproduction.

Besides the male competition signal, it has yet to be established whether other forms of signals could disrupt BPH courtship and mating behavior more effectively. In this paper, we developed a system for collecting, monitoring, and playing back BPH vibration signals. Artificial digital signals of different frequencies and forms were generated using Python and Adobe Audition according to the sensitive frequencies of BPH courtship calls and male competition calls. The playback and monitoring experiments were conducted to determine the most effective courtship disruptive signals.

Materials and methods

BPH rearing

All tested BPH subjects were obtained from the insect rearing room of the China National Rice Research Institute. The BPHs were reared on fresh rice seedlings at 26 ± 1°C and 70 ± 10% humidity under a 16/8 h light/dark photoperiod. Newly-emerged virgin BPHs (<24 h) were transferred to separate cages to avoid mating.

Development of BPH vibration signal recording, monitoring, and playback system

To collect, monitor, and play back the vibration signals of BPHs, we developed a novel system with magnetoelectric

converters, low-noise amplifier circuits, transducers, and other components. The setup is illustrated in Figure 2.

The magnetoelectric converter (model AT-3600 L) consists of a magnet, coil, and stylus. It converts the weak BPH vibration signal into a current output. The low-noise amplifier circuit utilizes the Texas Instruments bipolar high-performance audio operational amplifier (OPA1612), which has excellent noise performance and ultra-low distortion. To filter out noise from the BPH vibration signal and the signals in the non-target frequency band, we designed a Chebyshev sixth-order active band-pass filter. The system is powered by a lithium battery and the DC power circuit uses two integrated circuit chips: TPS563201 and LM27762. These devices support such functions as soft start, current limit, and thermal protection, as well as other functions. This power circuit has low noise and high-efficiency characteristics, and the 4,800 mah battery can be used continuously for more than 5 days. The transducer incorporates a 0.25 W flat-panel-driven full-frequency speaker. The signal collected through playback has no obvious distortion and only part of the frequency component is lost due to the mechanical filtering characteristics of rice plants. The system integrates power supply, amplification, and filter circuits with single-ended shielding and an optimized printed circuit board design to reduce noise during recording and monitoring.

Recording of BPH courtship vibration signals

A healthy rice plant was planted in a small pot and the stems and leaves of the rice plant above 15 cm were removed. The BPH test subject was placed in a circular observation box that was positioned about 10 cm above the rice stem base. At the top of the main rice stem, an iron nail was inserted as the contact point of the magnetoelectric converter. The converter of the system was attached to an iron stand with an adjustable shaft. Additionally, the position of the counterweight was modified so that the converter rested lightly on the contact surface directly above the nail. The output of the system was connected to a recording pen for recording and to headphones for monitoring. In this study, a total of 30 male BPHs and 30 female BPHs were used to record signals.

Generation of artificial disruptive signals

To disrupt the recognition of BPH courtship signals and male localization, we generated seven distinct disruptive signals of three different types. From the spectrum of the BPH courtship vibration signal in Section 3.1, the frequency of the female courtship signal is significantly in the range of 150 Hz to 300 Hz. Therefore, we selected the frequencies of 150 Hz, 225 Hz, and 300 Hz as the basis for the generation of artificial disruptive signals. Type 1 was the pure tone (PT) of different frequencies within the range of the female courtship signal. There were three kinds of PTs with frequencies of 150 Hz, 225 Hz, and 300 Hz. Type 2 was the continuous pulse signal (CPS), which had the same form as the female courtship signal. The three kinds of CPS had a constant pulse rate (PR) of 22 Hz and three different main vibration frequencies (MVF) of 150 Hz, 225 Hz, and 300 Hz. The Type 3 signal was white Gaussian noise (WGN) with 0 dBW power.

All disruptive signals were synthesized using the Python NumPy library (NumPy, 2002), in which the CPS was obtained by filling a standard sine wave with a triangular pulse at a frequency of 22 Hz. The oscillograms and spectrograms of the seven artificial disruptive signals are presented in Figure 3.

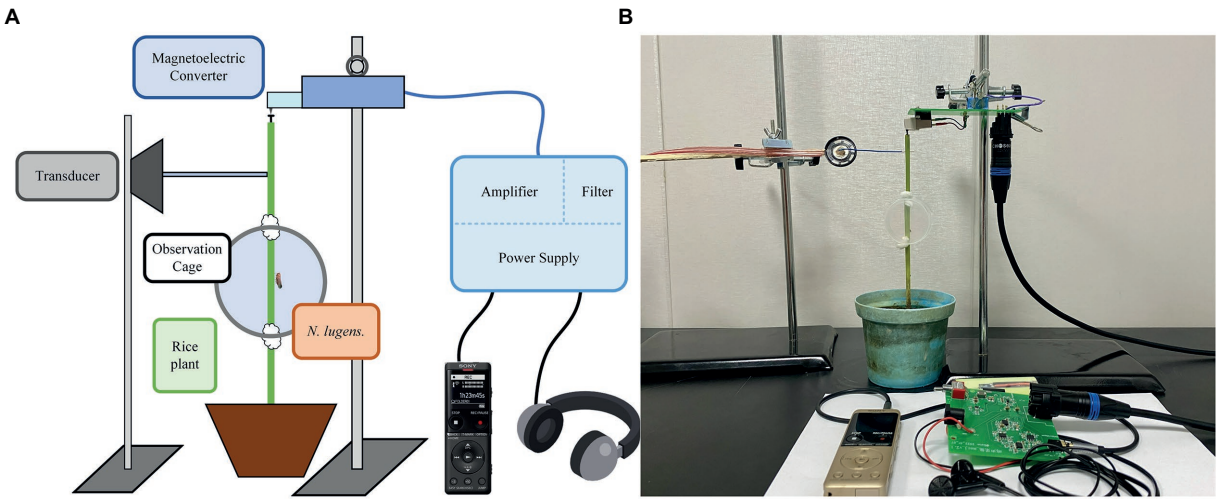


FIGURE 2
Vibration signal recording, monitoring, and playback system of BPH: **(A)** System framework diagram; **(B)** Physical image of the system.

Playback of artificial disruptive signals

The experimental conditions were the same as the BPH rearing conditions. Healthy and insect-free rice seedlings were selected and pruned according to the experimental requirements. Well-developed male BPHs that had not mated after emergence for 2–3 days were placed into the circular observation box. Next, the box and BPH were left alone for more than 3 h. The playback device was placed at a height of 2 cm above the box and the signals were played back using a Sony recording device (model ICD-UX575F). To avoid the influence of intensity difference in the courtship and disruptive signals on BPH courtship responses, the pre-recorded courtship signal, and disruptive signal were mixed according to a uniform intensity.

To ensure that each tested male BPH was active, a female courtship signal was played to each BPH. If the tested BPH replied to the played signal within 10 s twice in a row, it was considered to be in the active reply period (ARP), which means that the tested BPH reacted strongly to the courtship signal and sent signals in response. Experiments were carried out on BPH that was in the ARP, and a combination of two control tests (courtship signal) and two disruptive tests (mixed-signal involving the courtship and disruptive signals) was used to test each disruptive signal in a loop. Each signal was played one syllable at a time by referring to segments of the courtship signals of BPHs. The responses of the BPHs were observed within 45 s after playback. If the BPHs replied, the delay of the reply was recorded. The main purpose of recording whether the males replied was to establish whether males could distinguish female courtship signals mixed with various kinds of disruptive signals. The reply delay was the time it took for the males to recognize and reply to the courtship signal.

After each signal was played back, we waited 60 s to confirm whether the tested BPH replied. If it did not reply, we waited a further 60 s before conducting the next test. If there was still no reply after the courtship signal was played a second time, we performed an additional test of the courtship signal. If the tested BPH still did not reply, we concluded that the tested BPH was temporarily not in the ARP, and this data was not recorded in the experimental results. Otherwise, the experiments continued. Besides, various behaviors of the tested BPHs, including movement, searching, localization, etc., were observed and recorded throughout the experiments. Each group of playback experiments is independent of the other, and the BPH subjects were discarded at the end of the experiment, meaning the number of tests per set was equal to the number of subjects.

Evaluation of disruption effect on courtship behavior

To test whether the disruptive signals had a significant effect on the replies of males, a nonparametric test was performed on the data of reply delays. Besides, pairwise tests were conducted for paired data. Reply delays for the disruption-free signals and the seven kinds of artificially synthesized disruptive signals were compared using the Kruskal–Wallis test followed by Dunn's pairwise comparison test (Kruskal and Wallis, 1952; Dunn, 1964). In this study, the statistical analysis was conducted using the Kruskal–Wallis H test in SPSS Statistics.

Results and analysis

Spectra of BPH courtship vibration signals

In the experiment, we recorded 60 female courtship signals, 66 male courtship signals, and 39 male competition signals. Analysis of the BPH courtship signals indicates that the female signals exhibit a continuous pulse signal. Besides, by using the fast Fourier transform to analyze the frequency spectrum of the signal, we determined that its main vibration frequency is 236 ± 43 Hz and the pulse rate is 23 ± 2 Hz. The male signal is more complex and can be divided into a courtship signal and a competition signal. The courtship signal of males can generally be divided into three parts. The first part is three to ten irregular pulses (a), followed by a continuous pulse signal with a main vibration frequency of 255 ± 24 Hz (b), while the third part is zero to five wide pulses (c). The male BPH competition signal consists of continuous pulses (a) and then two to four short pulses (b), and the main frequency of the continuous pulse segment is 281 ± 46 Hz. The oscillograms and spectrograms of the three signals are displayed in Figure 4.

Courtship disruption effect of disruptive signals

Playback experiments with disruption-free signals and seven kinds of artificially synthesized disruptive signals were carried out. The results of the reply rate and reply delay time are shown in Table 1.

Table 1 reveals that the male BPHs replied to female courtship signals in 95.6% of the control experiments. Besides, using Gaussian white noise resulted in a reply rate that was close to the control group, at 83.7%. In contrast, under the pure tone disruptive signal of 225 Hz, the reply rate of males was only 33.3%. The reply rate obtained with other disruptive signals ranged from 42.9 to 66.7%. As illustrated in Figure 5, there were significant differences in reply delay for BPHs that replied (Kruskal–Wallis test, $H = 19.550$, df = 7, $p = 0.007^{**}$). Also, the results of the pairwise comparisons showed significant differences between the three groups of PT experiments and the control experiments (Dunn's pairwise comparison test, Pcontrol-150 Hz = 0.009**, Pcontrol-225 Hz = 0.007**, Pcontrol-300 Hz = 0.005**). However, there were no significant differences for the other disruptive signals. The disruptive signals caused a substantial decrease in the reply rate of the tested BPHs and a considerable increase in the reply delay of

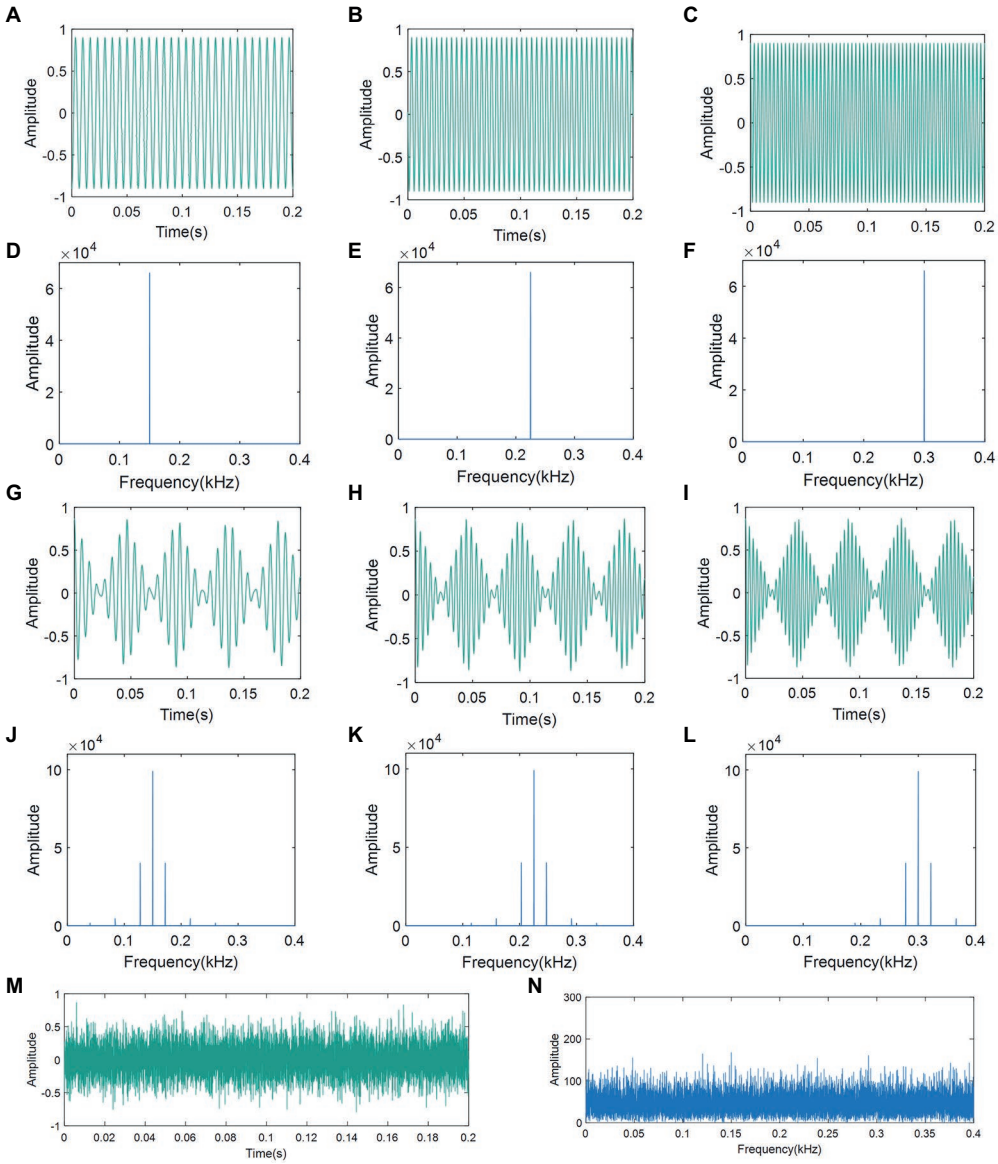


FIGURE 3
Oscillograms and spectrograms of artificial disruptive signals: **(A)** Oscillogram of PT (150Hz); **(B)** Oscillogram of PT (225Hz); **(C)** Oscillogram of PT (300Hz); **(D)** Spectra of PT (150Hz); **(E)** Spectra of PT (225Hz); **(F)** Spectra of PT (300Hz); **(G)** Oscillogram of CPS (22-150Hz); **(H)** Oscillogram of CPS (22-225Hz); **(I)** Oscillogram of CPS (22-300Hz); **(J)** Spectra of CPS (22-150Hz); **(K)** Spectra of CPS (22-225Hz); **(L)** Spectra of CPS (22-300Hz); **(M)** Oscillogram of WGN; and **(N)** Spectra of WGN. PT: pure tone; CPS: continuous pulse signal; WGN: white Gaussian noise.

the BPHs that replied. This indicates that playing the disruptive signal during the male recognition process of female courtship signals effectively disrupts the replies and localization of males, with a disruption rate of more than 50%. Also, the males that reply are less efficient in courtship due to the influence of the disruptive signals. Among the various disruptive signals tested, the disruptive effect of PT (225 Hz) was the best, followed by PT (300 Hz). The other synthetic disruptive signals also had a certain disruptive influence, while the WGN signal had almost no effect. Therefore, we presume that the BPH courtship signal recognition process is particularly sensitive to frequency. In cases when the CPS has the same intensity as the PT, the width of the CPS spectrum is greater. This means that the intensity of the CPS at the same frequency point is slightly weaker and the disruptive effect is lower than PT.

Behavioral response

In the control experiments, the tested male BPHs exhibited stimulation after receiving the female courtship signals. There were two general modes of the tested BPHs in the behavioral observations. In Mode 1, the BPHs searched and localized during playback of the courtship signal and replied soon after the signal stopped. In Mode 2, the BPHs did not exhibit obvious searching or

10.3389/fpls.2022.897475

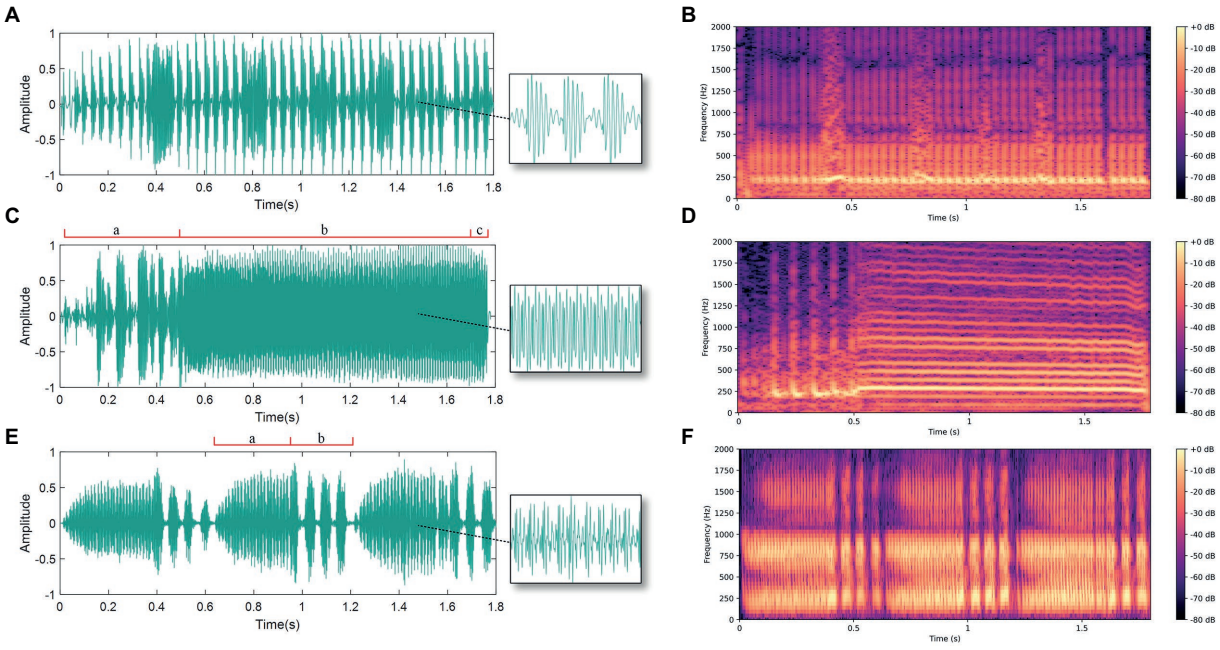


FIGURE 4
Oscillograms and spectrograms of male and female vibration signals: **(A)** Oscillogram of female courtship signal; **(B)** Spectrogram of female courtship signal; **(C)** Oscillogram of male courtship signal (divided into parts a, b and c); **(D)** Spectrogram of male courtship signal; **(E)** Oscillogram of male competition signal (divided into parts a and b); **(F)** Spectrogram of male competition signal.

TABLE 1 Screening results of BPH courtship-specific disruptive signals.

Disruptive signals		No. of tests	No. of replies	Reply rate	Mean (±SD) reply delay (s)
Control	None	45	43	95.6%	5.3 ± 3.3
Artificial Disruptive Signals	PT (150 Hz)	39	22	56.4%	8.0 ± 5.2
	PT (225 Hz)	42	17	33.3%	9.1 ± 5.6
	PT (300 Hz)	42	18	42.9%	9.3 ± 5.9
	CPS (PR: 22 Hz; MVF: 150 Hz)	42	28	66.7%	5.8 ± 3.2
	CPS (PR: 22 Hz; MVF: 225 Hz)	40	21	52.5%	7.2 ± 5.0
	CPS (PR: 22 Hz; MVF: 300 Hz)	40	22	55.0%	7.4 ± 6.0
	WGN	43	36	83.7%	5.0 ± 3.1

PT, pure tone; CPS, continuous pulse signal; WGN, white Gaussian noise; PR, pulse rate; and MVF, main vibration frequency.

localizing behavior during the playback of the courtship signal, but they still replied to the playback signal after the signal stopped. In a large number of experiments, both modes were presented and the behavioral mode of a single tested BPH switched from time to time.

In the experiments using disruptive signals, the aroused behavior of the tested BPHs was lower than in the control experiments. There was no reply to the playback signal from 62.2% of the tested BPHs, and behavioral observations indicated that the tested BPHs did not show any obvious movement during playback or after playback stopped. The BPHs that replied were mostly of the second behavioral mode mentioned above, meaning that active searching and locating behavior was very rare. These observations demonstrated that the disruptive signals had a significant effect on the courtship behavior of BPHs.

Discussion

The rice pest BPH does not have sex pheromones. Instead, it vibrates its abdomen at specific frequencies for courtship, localization, and eventually mating and egg-laying (Zhang and Chen, 1987a). To date, chemical control is still the primary method used to regulate the population of BPHs. This has caused increased BPH resistance to pesticides, rice pesticide residue, and serious environmental pollution (Wu, 2018). Thus, it is beneficial to use disruptive signals that upset the courtship of male and female BPHs, thereby delaying mating and reducing the mating rate. The bioacoustic method for disrupting the mating of *S. titanus* is currently under development and it has presented proof of concept under small-scale field conditions

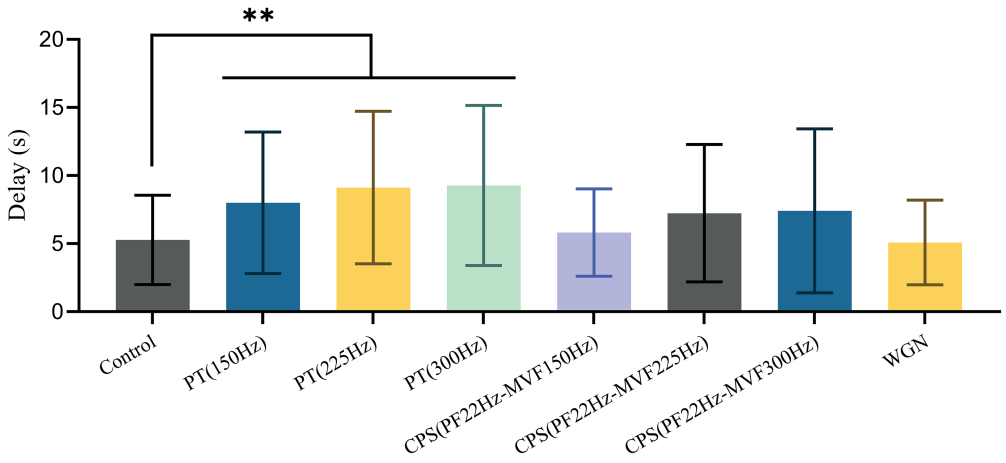


FIGURE 5
Mean±SD of male reply delay from the female courtship signal with or without disruptive signals. Differences from the Kruskal–Wallis test followed by Dunn's pairwise comparisons are denoted as follows: ** p <0.01.

(Polajnar et al., 2016). While current studies on the vibration control of BPHs are limited to the disruptive effect of male competition signals (Fu et al., 1997, 1999), this paper aims to extract more effective disruptive signals for disturbing courtship responses.

Currently, costly equipment such as the laser vibrometer is commonly used in vibration studies of *S. titanus* (Mazzoni et al., 2009). In this paper, a system for collecting, monitoring, and playback of BPH vibration signals was developed using magnetoelectric converters, low-noise amplifier circuits, transducers, and other components. The system we devised achieved higher functionality with lower costs. Thus, it has the potential to be applied to more species of planthoppers and leafhoppers, thereby facilitating more insect vibration studies.

Vibration courtship disruption involves the use of specific disruptive signals to precisely upset the communication of BPHs which are in the courtship stage. It blocks or disrupts the normal courtship behavior between males and females and effectively prevents or controls overall mating behavior. BPH courtship signals stimulate sexual arousal in the opposite sex and are so specific that BPHs are insensitive to other types of signals. Therefore, the development of disruptive signals using the characteristics of courtship signals may affect the perception process of courtship signals. This is different from the method devised by Fu et al. (1997, 1999), which utilized male competition signals. The disruption method of the competition signal does have some influence, but the mutually competitive behavior of males is quite extensive in the BPH population and the disruption effect is limited. In contrast, the specific disruptive signal obtained through screening emphasizes the perception process. Therefore, it has better potential and wider applicability. Furthermore, analysis results of male and female courtship signals show that although the signals differ greatly in the time domain, the frequency range of the main vibrations is relatively stable. Therefore, normal courtship communication can be disrupted by playing back disruptive signals in the same frequency range during the courtship process between male and female BPHs. However, the BPH courtship and localization process is very conservative, with males generally initiating calls and then locating the female based on the female's response (Zhang, 1997). Also, the male courtship signals are slightly more complex in the time domain, with different main vibration frequencies for male and female courtship signals. Therefore, we focused on the male recognition process of female signals to design experiments that screened disruptive signals for upsetting the courtship process of BPH males and females.

By using artificially synthesized disruptive signals with different frequency combinations for playback and monitoring experiments, a pure tone signal of 225 Hz was extracted. This signal affected male BPH recognition of the female's courtship call. Experiments revealed that the male BPH reply rate decreased from the original 95.6 to 33.3%. Furthermore, the mean reply delay increased from 5.3 s to 9.1 s. This confirmed that the mean reply delay of males under disruption was significantly different from the control group. Therefore, it is feasible to use specific disruptive signals to disrupt the courtship behavior of BPHs. We hypothesize that the effectiveness of the 225 Hz pure tone is due to its disruption of the spectral structure of the female courtship signal, making it difficult for males to recognize or become sexually aroused.

By applying the mating disruption technique for *S. titanus*, the mating inhibition rate was about 90% (Polajnar et al., 2016). Due to its satisfactory experimental effects, this method has tangible potential for controlling *S. titanus* reproduction. The experiments in this paper only focused on behavior at the courtship stage, but our results confirmed the hypothesis that disruption with specific disruptive signals efficiently disturbs BPH courtship behavior. Furthermore, our results provide support for subsequent research on BPH control methods.

Admittedly, there are limitations to the findings of our study. The effects of the duration and frequency of the disruptive signals on courtship behavior, as well as on mating and egg-laying, still need to be tested experimentally. Additionally, the experiments in this paper were conducted indoors. However, if the results are applied to paddy fields, there will be a great deal of external interference noise, such as the shaking of the rice plants due to wind, noise from the surrounding environment, etc. For *S. titanus*, standards in viticulture provide delivery means of vibrational energy to target surfaces (Polajnar et al., 2016). However, for BPHs, the choice of delivery medium is still an issue. If the experimental technique is applied to paddy fields, substantial research efforts are still required to determine what equipment and what medium can be used to effectively deliver disruptive signals to BPHs on rice stems. In the next stage of our research, we will observe the effect of simulated external noise on BPH courtship and further investigate the effect of disruptive signals on BPH mating. Moreover, we will enhance the playback equipment to address certain issues such as avoiding environmental pollution, reducing costs, and expanding the effective coverage range. These will be very challenging but interesting tasks.

Data availability statement

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

Author contributions

ZF, ZY, and QY came up with the idea and wrote and revised the manuscript. ZF, QW, and BY purposed the experimental methods. YG, JL, YD, and JB analyzed the experimental data. All authors contributed to the article and approved the submitted version.

Funding

This study is supported by the Key R&D Program of Zhejiang (No. 2022C02004).

Acknowledgments

The authors would like to thank China National Rice Research Institute for providing the experiment materials and environment.

Conflict of interest

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

Publisher's note

References

Cokl, A., and Virant-Doberlet, M. (2003). Communication with substrate-borne signals in small plant-dwelling insects. *Annu. Rev. Entomol.* 48, 29–50. doi: 10.1146/annurev.ento.48.091801.112605

Dunn, O. J. (1964). Multiple comparisons using rank sums. *Technometrics* 6, 241–252. doi: 10.1080/00401706.1964.10490181

Dyck, V. A., and Thomas, B. (1979). "The brown Planthopper problem," in *Brown Planthopper: Threat to Rice Production in Asia*. ed. M. Pathak (Philippines: International Rice Research Institute), 3–17.

Eriksson, A., Anfora, G., Lucchi, A., Lanzo, F., Virant-Doberlet, M., and Mazzoni, V. (2012). Exploitation of insect vibrational signals reveals a new method of Pest management. *PLoS One* 7:e32954. doi: 10.1371/journal.pone.0032954

Fu, Q., Chen, W., Zhang, Z., and Tang, X. (1997). The second male vibrational signal of Brown Planthopper *Nilaparvata lugens* (Stål) and its significance in competitive reproductive behavior. *Acta Entomol. Sin.* 40, 254–260. doi: 10.16380/j.kcxb.1997.03.004

Fu, Q., Chen, W., Zhang, Z., and Tang, X. (1999). Inhibition of the second male vibrational signal on reproduction of the Brown Planthopper, *Nilaparvata lugens* (Stål). *Acta Entomol. Sin.* 42, 3–8. doi: 10.16380/j.kcxb.1999.01.001

Ichikawa, T., Sakuma, M., and Ishii, S. (1975). Substrate vibrations: mating signal of three species of Planthoppers which attack the Rice Plant: Homoptera: Delphacidae. *Appl. Entomol. Zool.* 10, 162–171. doi: 10.1303/aez.10.162

Kruskal, W. H., and Wallis, W. A. (1952). Use of ranks in one-criterion variance analysis. *J. Am. Stat. Assoc.* 47, 583–621. doi: 10.1080/01621459.1952.10483441

Matsumura, M., and Sanada-Morimura, S. (2010). Recent status of insecticide resistance in Asian Rice Planthoppers. *Jarq-Japan Agricult. Res. Q.* 44, 225–230. doi: 10.6090/jarq.44.225

Mazzoni, V., Lucchi, A., Cokl, A., Presern, J., and Virant-Doberlet, M. (2009). Disruption of the reproductive behaviour of Scaphoideus titanus by playback of vibrational signals. *Entomol. Exp. Appl.* 133, 174–185. doi: 10.1111/j.1570-7458.2009.00911.x

NumPy (2002). Available at: https://github.com/numpy/numpy (Accessed Jan 15, 2021).

Ossiannilsson, F. (1949). "Insect drummers: a study on the morphology and function of the sound-producing organ of Swedish Homoptera Auchenorrhyncha, with notes on their sound-production".

Polajnar, J., Eriksson, A., Virant-Doberlet, M., Lucchi, A., and Mazzoni, V. (2016). "Developing a bioacoustic method for mating disruption of a leafhopper Pest in grapevine," in *Advances in Insect Control and Resistance Management*. eds. A. R. Horowitz and I. Ishaaya (Cham: Springer International Publishing), 165–190.

Saxena, R. C., Zhang, Z. T., and Boncodin, M. (1993). Neem oil affects courtship signals and mating behavior of brown planthopper *Nilaparvata lugens* (Stål; Hom., Delphacidae) females. *J. Appl. Entomol.* 116, 127–132. doi: 10.1111/j.1439-0418.1993.tb01178.x

Wu, K. (2018). Development direction of crop Pest control science and Technology in China. *J. Agri.* 8, 35–38.

Zhang, Z. (1997). Overview of insect Acousitc communication and study on *Nilaparvata lugens* vibration signals. *Life Sci.* 9, 227–230.

Zhang, Z., and Chen, L. (1987a). Collection and analysis of vibration signals in three Planthoppers and leafhoppers. *Chin. Sci. Bull.* 32, 1583–1586. doi: 10.1360/csb1987-32-20-1583

Zhang, Z., and Chen, L. (1987b). A preliminary report on courtship signals os Planthoppers and leafhopper which attack the Rice Plant. *Chin. J. Rice Sci.* 1, 134–135. doi: 10.16819/j.1001-7216.1987.02.010

Zhang, Z., and Chen, L. (1991). Insect vibrational signal monitoring, recording and playback technology. *Chinese Bullet. Entomol.* 28, 170–172.

Zhang, Z., Yin, B., Chen, L., and Saxena, R. C. (1991). Courtship signals and mating behaviour of Rice Brown Planthopper, *Nilaparvata lugens* (Stål). *Acta Entomol. Sin.* 34, 257–265. doi: 10.16380/j.kcxb.1991.03.001

frontiers | Frontiers in Plant Science

TYPE Original Research
PUBLISHED 15 August 2022
DOI 10.3389/fpls.2022.931491

OPEN ACCESS

EDITED BY
Shanwen Zhang,
Xijing University,
China

REVIEWED BY
Therese McBeath,
Commonwealth Scientific and Industrial Research Organisation (CSIRO), Australia
Arturo Reyes,
Instituto Nacional de Investigaciones Agrícolas, Forestales y Pecuarias, Mexico

*CORRESPONDENCE
Rodrigo Filev Maia
r.filevmaia@deakin.edu.au

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

RECEIVED 29 April 2022
ACCEPTED 26 July 2022
PUBLISHED 15 August 2022

CITATION
Maia RF, Lurbe CB and Hornbuckle J (2022) Machine learning approach to estimate soil matric potential in the plant root zone based on remote sensing data. *Front. Plant Sci.* 13:931491. doi: 10.3389/fpls.2022.931491

Machine learning approach to estimate soil matric potential in the plant root zone based on remote sensing data

Rodrigo Filev Maia*, Carlos Ballester Lurbe and John Hornbuckle

Centre for Regional and Rural Futures, Deakin University, Hanwood, NSW, Australia

There is an increasing interest in using the Internet of Things (IoT) in the agriculture sector to acquire soil- and crop-related parameters that provide helpful information to manage farms more efficiently. One example of this technology is using IoT soil moisture sensors for scheduling irrigation. Soil moisture sensors are usually deployed in nodes. A more significant number of sensors/nodes is recommended in larger fields, such as those found in broadacre agriculture, to better account for soil heterogeneity. However, this comes at a higher and often limiting cost for farmers (purchase, labour costs from installation and removal, and maintenance). Methodologies that enable maintaining the monitoring capability/intensity with a reduced number of in-field sensors would be valuable for the sector and of great interest. In this study, sensor data analysis conducted across two irrigation seasons in three cotton fields from two cotton-growing areas of Australia, identified a relationship between soil matric potential and cumulative satellite-derived crop evapotranspiration (ET_{cn}) between irrigation events. A second-degree function represents this relationship, which is affected by the crop development stage, rainfall, irrigation events and the transition between saturated and non-saturated soil. Two machine learning models [a Dense Multilayer Perceptron (DMP) and Support Vector Regression (SVR) algorithms] were studied to explore these second-degree function properties and assess whether the models were capable of learning the pattern of the soil matric potential-ET_{cn} relation to estimate soil moisture from satellite-derived ET_c measurements. The algorithms performance evaluation in predicting soil matric potential applied the k-fold method in each farm individually and combining data from all fields and seasons. The latter approach made it possible to avoid the influence of farm consultants' decisions regarding when to irrigate the crop in the training process. Both algorithms accurately estimated soil matric potential for individual (up to 90% of predicted values within ±10kPa) and combined datasets (73% of predicted values within ±10kPa). The technique presented here can accurately monitor soil matric potential in the root zone of cotton plants with reduced in-field sensor equipment and offers promising applications for its use in irrigation-decision systems.

KEYWORDS

machine learning, remote sensing, soil matric potential, NDVI, evapotranspiration, irrigation

Introduction

Population growth in recent decades has boosted water demand worldwide, but water use and water consumption trends at current rates are unsustainable (FAO, 2017). Therefore, multidisciplinary research efforts in water management are required to achieve a sustainable future (Cosgrove and Loucks, 2015). This includes research on water management in agriculture, which accounts for ~70% of the freshwater use worldwide and is the most water-demanding of all economic sectors (Pimentel et al., 2004). Irrigation is the activity that requires most of the water resources available and influences a variety of biophysical processes in plants that directly relate to yield.

Optimizing irrigation scheduling decisions in agriculture is a challenge that needs to be met to manage water resources more efficiently and improve crop water productivity (Jägermeyr, 2020; Chaudhary and Srivastava, 2021). Technology development is critical in providing more accurate tools for monitoring soil and crop water status at different scales (Bittelli, 2011; Cahn and Johnson, 2017; Saad and Gamatié, 2020). A range of approaches are available for direct and indirect measurements of the soil and plant water status (see Jones (2006)). For irrigation scheduling purposes, soil moisture monitoring has traditionally been the methodology used, and it is generally preferred to other methods due to its suitability for irrigation automation. An evaluation of available soil moisture measurement technologies and their limitations can be found in Susha Lekshmi et al. (2014). Water content sensors provide helpful information to determine when to irrigate, but soil texture influences the measurements requiring site-specific calibration (Cahn and Johnson, 2017). Soil matric potential sensors indicate how readily water is accessible for plants (Jones, 2006), do not require soil-specific calibration and are generally preferred for water management in vegetable crops (Thompson et al., 2007; Cahn and Johnson, 2017).

One of the limitations to the wide use of soil matric potential sensors at the commercial scale is that they only provide point-source measurements, and thus, many sensors are needed in large-scale broadacre farming to accounting for soil spatial variability (Cahn and Johnson, 2017). To monitor large heterogeneous farms with sensors, Wireless Sensor Networks can be used to interconnect them and make data available online in near real-time. However, it considerably increases the cost of the monitoring system (acquisition, installation, maintenance, among others), which has been reported as one of the significant barriers to the wide adoption of these technologies by farmers (Blasch et al., 2022). Remote sensing techniques can estimate soil moisture and allow capturing the existing spatial variability in large areas. However, these techniques still need improvement and are not accurate enough to directly estimate soil moisture at a field scale suitable for irrigation scheduling (Shunlin Liang, 2020). Therefore, methodologies that could minimize the number of in-field sensors without losing soil moisture monitoring capability/intensity would greatly value water managers.

While not yet ready to accurately directly monitor soil moisture at the field scale for irrigation scheduling, remote sensing-based approaches are helpful for the estimation of crop evapotranspiration through vegetation indices gathered from satellite, airborne and drone-based platforms following the FAO56 method (Pereira et al., 2015; Pôças et al., 2020). In this approach, crop evapotranspiration (ET_c) is estimated by multiplying the reference evapotranspiration (ET_o) obtained from data collected at nearby weather stations by a site-specific crop coefficient that is obtained from vegetation indices that can be monitored at high temporal and spatial resolution. The Normalized Difference Vegetation Index (NDVI) is the most widespread vegetation index and is linearly related to the crop coefficient (Trout and Johnson, 2007). This ET_c/NDVI approach enables water managers to monitor crop water requirements at individual sites throughout the growing season. However, the rate at which moisture is depleted from the soil is related to the crop evapotranspiration, which varies with the crop phenological stage, and thus, their relationship is time dependent. This time-dependent relationship hinders the possibility of estimating or accurately predicting soil matric potential from evapotranspiration measurements by conventional data processing techniques. In this work, we studied this relationship when soil matric potential is used to monitor soil moisture in gravity surface irrigated systems. The hypothesis was that machine learning models can learn the interaction between soil-, crop-, and weather-related parameters to estimate soil matric potential in the root zone from remotely sensed evapotranspiration measurements.

With the adoption of Information and Communication Technology in agriculture and the substantial volume of data generated, data-driven machine learning techniques that can organize data from different sources and the power to learn from them become essential (Alzubi et al., 2018; Benos et al., 2021). Machine learning techniques have been applied in agriculture in various farming practices, with those related to crop management activities (disease detection, yield prediction, among others) receiving most of the attention. Although growers see potential in using the Internet of Things, remote sensing, and machine learning (Agriculture 4.0) for having better decision-making processes, particularly in irrigation, substantially less work has been undertaken on water management activities (Benos et al., 2021). Within the studies focused on water management, machine learning techniques have been applied to estimate groundwater reservoirs, soil moisture (Paloscia et al., 2013; Coopersmith et al., 2016; Prasad et al., 2018; Singh et al., 2019; Babaeian et al., 2021; Greifeneder et al., 2021; Grillakis et al., 2021; Orth, 2021; Sungmin and Rene, 2021), evapotranspiration (Ponraj and Vigneswaran, 2020), and provide irrigation control (González-Briones et al., 2019; Kondaveti et al., 2019; Murthy et al., 2019; Akshay and Ramesh, 2020; Campoverde et al., 2021; Ikidid et al., 2021; Perea et al., 2021; Bhoi et al., 2021a), among other applications (Liakos et al., 2018; Cardoso et al., 2020; Perea et al., 2021; Bhoi et al., 2021b). The machine learning techniques applied in these studies are shown in Table 1, following the classification suggested in

TABLE 1 Algorithms used for a range of applications in agriculture using remote sensing (R) and sensors data (S).

Application	BM	SVM	ANN	Regression	DT	MAS	Genetic
Water estimation	–	S	S	S	–	–	–
Soil moisture	R, S	–	R, S	R, S	R, S	S	–
Evapotranspiration	–	–	R, S	S	S	–	R, S
Irrigation control	S	S	S	–	S	–	–
Rainfall prediction	–	S	–	–	–	–	–
Total	4	4	9	5	9	2	1

BM, Bayesian models; SVM, support vector machine; ANN, artificial neural network; DT, decision tree; MAS, multi-agent system.

(Liakos et al., 2018) and considering two additional categories: Multi-Agent System (MAS) and Genetic Algorithm. The algorithms applied to estimate soil moisture are Bayesian models, Artificial Neural Networks (ANN), Regression models, Decision Tree models and MAS. Several research papers rely on neural network algorithms to classify or estimate crop parameters. Support Vector Regression (SVR, also called SVM in most papers as shown in Table 1) and Decision Tree-based algorithms are also primarily used for the same purposes.

This work proposes the original approach of using the relationship between soil matric potential and the cumulative evapotranspiration between irrigation events expressed in kPa/mm aiming to (i) explore the proposed relation over the cotton growing season and (ii) assess the feasibility of estimating soil matric potential in the cotton root zone (0.20 m below ground) from remotely sensed evapotranspiration by using machine learning models. The kPa/mm relation may represent the dynamics of crop water use during the season. Supported vector models and ANN were the machine learning models applied because of their capability to process time-dependent parameters. The models' performance in estimating soil matric potential at bay level was evaluated and compared following two approaches: (i) when models were trained with data for each farm and growing season, and (ii) when models were trained with data from all the sites and seasons combined. The second approach was implemented to avoid any influence water managers' decision practices could have on the algorithm responses.

The study contributes to the research on implementing machine learning techniques in irrigation water management that are scarce in the literature compared to crop management activities. It presents an approach to cotton producers of the main cotton-growing areas of Australia that would allow them to monitor soil matric potential with a reduced number of in-field sensors and potentially optimize on-farm water management in these systems.

Materials and methods

Site locations and characteristics

The study was conducted during two cotton-growing seasons (2019/20 and 2020/21) with data collected from three commercial irrigated cotton farms located in the Murrumbidgee Valley (sites A and B) and Moree Plains Shire (site C) in the south and north of NSW, Australia, respectively (Figure 1). During a typical cotton growing season, these farms have approximately 500–1,500 irrigated hectares depending on irrigation allocations. Irrigation fields typically have bays ranging from 8–30 ha, so hundreds of irrigated bays may need to be managed for irrigation water applications during the irrigation season.

In the 2019/20 growing season, soil matric potential was monitored at sites A and B, while in the 2020/21 season, it was monitored at sites A and C. The Murrumbidgee area (sites A and B) climate is semi-arid, while it is humid subtropical in Moree in the north of NSW. The weather conditions differed between growing seasons and between cotton producing areas. Total rainfall and reference evapotranspiration (ET_o) for sites A and B in the 2019/20 cotton growing season (from mid-October to April) were 152 and 1,161 mm, respectively. In the 2020/21 growing season, total rainfall and ET_o were 305 and 1,224 mm at site A and 582 and 1,245 mm, respectively, at site C.

Three bays were monitored at sites A (14.6 ha in total) and B (21.0 ha), while five bays were monitored at site C (172.8 ha) in the Moree Plains cotton-producing region where the standard practice is to produce in larger bays (up to ~38 ha in farm C). In all sites, cotton was furrow irrigated employing a bank-less channel irrigation system (Grabham, 2012). Irrigation was scheduled based on the farm consultants' decision except for site A during the 2020/21 season, where irrigation was triggered based on soil matric potential thresholds and recommendations obtained from the cloud-based IRRISENS platform (Filev Maia et al., 2020). At this site, the grower was notified by text message when soil matric potential at 0.20 m depth at any of the three bays monitored was lower than −40 kPa to be able to order water to the irrigation company and organize the irrigation event 48 h in advance.

Measured and estimated parameters

Soil matric potential and crop evapotranspiration (ET_c) were measured and estimated at each site to explore their relationship between irrigation events over the growing seasons. These parameters, rainfall and the growing degree days (GDD), were

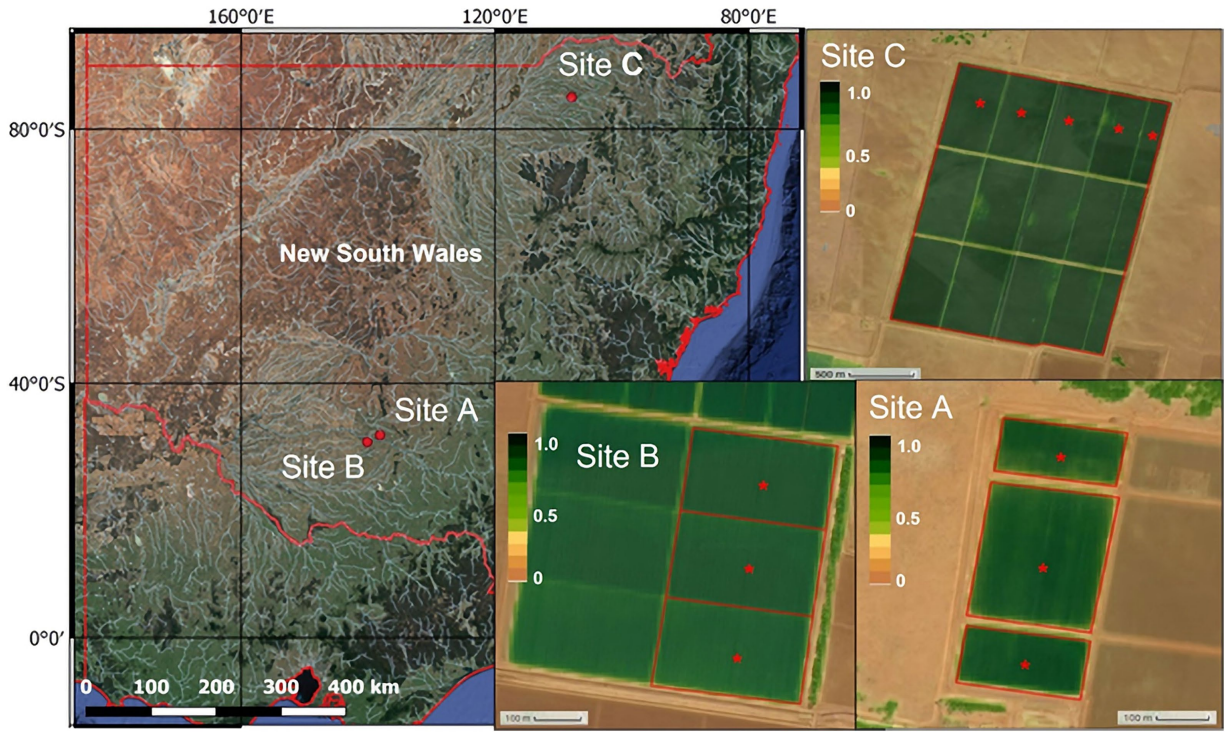


FIGURE 1
Location of the farms in the Murrumbidgee Valley (sites A and B) and Moree Plains Shire (site C) in NSW, Australia, where soil matric potential was monitored for this study. The red asterisks shown in each NDVI (normalized difference vegetation index) map of each site indicate the monitoring stations. Each station was composed of two matric potential sensors and one temperature sensor buried at 0.20m depth connected to a WiField datalogger.

used as inputs in the machine learning models described in the following subsection.

A WiField logger (Brinkhoff et al., 2017) with two watermark sensors (Model 200SS, Irrometer Company Inc., CA, United States) and one 1-wire temperature shielded sensor (model DS18B20) was used to continuously monitor soil matric potential at each bay. Watermark sensors were installed at 0.20 m below ground as in Kang et al. (2012) and Ballester et al. (2021), where soil matric potential measurements are essential to trigger irrigation events (Brinkhoff et al., 2017). Soil matric potential was calculated using the resistance of each watermark and the soil temperature based on the following equations (Irrometer, 2021):

$$\begin{cases} st = 0, \quad r < 550\Omega \\ st = \left[\left(r * 10^{-3}\right) * 23.156 - 12.736\right] - \left[1 + 18 * 10^{-3} * \left(t - 24\right)\right], \\ 550\,\Omega \le r < 1\,k\Omega \\ st = \dfrac{-3.213 * \left(r * 10^{-3}\right) - 4.093}{1 - 0.009733 * \left(r * 10^{-3}\right) - 0.01205 * t}, \\ 1k\Omega \le r < 8\,k\Omega \end{cases}$$

Figure 2 presents the seasonal evolution of the soil matric potential for each season and site. When soil matric potential readings are observed with more detail between irrigation events, there is an inflexion point that indicates the transition from a saturated (soil matric potential ≥ -10 kPa) to a non-saturated sate (Figure 3). Thus, soil matric potential evolution between irrigation events can be modelled by a second-degree function. The second-degree equation coefficients indicate that the soil matric potential has approximately a linear behavior after the inflexion point. The same behavior was observed at different moments in the season (different crop phenological stages) with weather conditions influencing the soil matric potential decrease rate.

The ET_c was estimated following the FAO56 approach (Allen et al., 1998), in which water requirements are obtained as the product between reference evapotranspiration (ET_o) and a crop coefficient ($ET_c = ET_o \times K_c$). ET_o (using alfalfa as the reference crop) was estimated by the Penman-Monteith equation using weather information obtained from the nearest meteorological stations to each site. The crop coefficient was estimated from satellite-based NDVI images using the relationship reported in Trout and Johnson (2007) ($K_c = 1.37$ NDVI – 0.086). NDVI images of each site were obtained using the Google Earth Engine API to access the Sentinel-2 top of atmosphere reflectance data collection. The extracted images were processed to eliminate those with more than 5% of pixels with clouds. The remaining

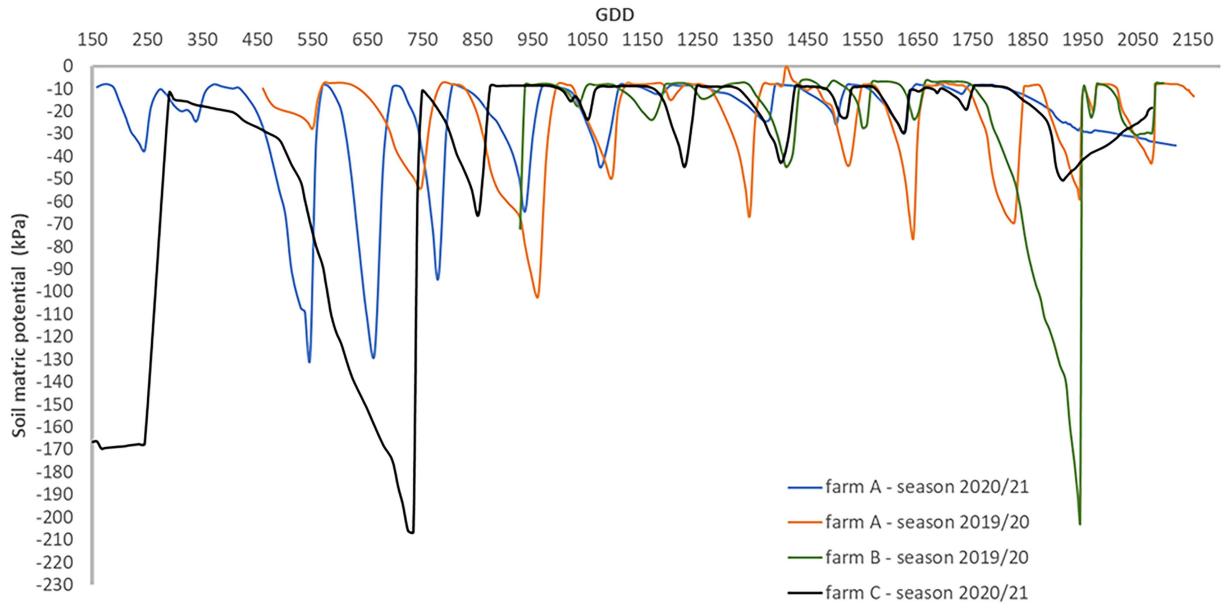


FIGURE 2
Seasonal soil matric potential evolution measured at 0.20m depth for each farm and cotton growing season.

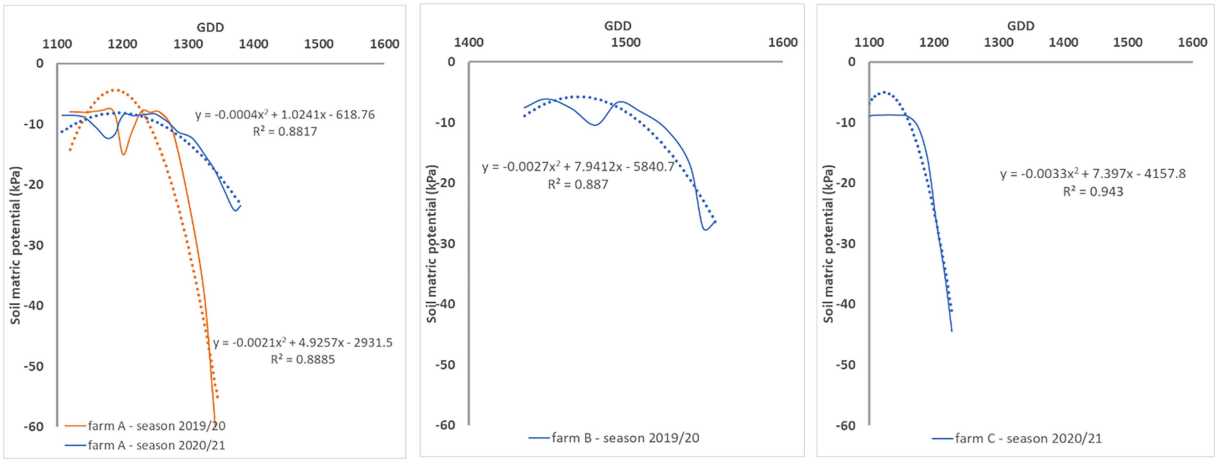


FIGURE 3
Illustration of the soil matric potential evolution between irrigation events at each farm for different crop phenological stages. The least square function and coefficient of determination (R^2) is shown for each irrigation event.

images were used to calculate the NDVI according to. As Sentinel-2 satellites do not provide daily data in the monitored areas, NDVI was estimated using linear regression for those dates with no images available.

Daily ET_c was used to calculate the cumulative evapotranspiration between irrigation events as follows:

$$ET_{cn} = \sum_{d=0}^{n} ET_c(d)$$

where d is days since the last irrigation, its value is zero on the day that irrigation occurs, and maximum n on the day before the next irrigation. The cumulative ET_c has the same cumulative behaviour as the soil matric potential.

The GDD index was another input parameter used in the models to account for the effect of temperature on crop development in each cotton-producing region (McMahon and Low, 1972). This index is computed from daily maximum and minimum air temperature data and a temperature threshold (the base temperature) at which crop growth stops. That for cotton is 12°C as follows:

TABLE 2 Potential input features used in this study to estimate soil matric potential.

Features	Source	Dimension/frequency
Solar radiation	Weather station	Daily value
GDD	Weather station	Daily value
NDVI	Remote sensing	Daily estimated
ETc_n	calculated	Cumulative evapotranspiration between irrigation events (daily)
$Rain_n$	Weather station	Cumulative rainfall in mm between irrigation events (daily)
Soil matric potential	Sensor/estimated	Daily average

GDD, growing degree days; NDVI, normalized different vegetation index; ET_c, cumulative crop evapotranspiration.

$$GDD = \sum_{k=0}^{n} \frac{(Tmax_k - 12) + (Tmin_k)}{2}$$

Machine learning proposed approach

Data selection and pre-processing

In all monitored sites there were three sources of data: WiField loggers, Sentinel-2 data collections, and weather stations installed in each site or close to it. Table 2 depicts the machine learning input data collected in the field. Data cleaning and organization were the same in all seasons according to the following criteria:

- soil matric potential readings range between 0 kPa and −200 kPa – readings lower than this value do not represent correct values;
- soil matric potential at each bay was the average of two sensors;
- weather data was collected from on-site weather stations (sites A and C) or nearby (site B);
- remote sensing data extraction occurred every 5 days when Sentinel-2 covered the farm area at least once;
- each satellite image and data extraction considered only points from the polygon representing each bay of the evaluated paddock, and the cloud coverage in such polygon must be less than 5% of the pixels in the image.

All data were organized sequentially according to the GDD to enable comparisons between crops from different locations and seasons. Two datasets were created to evaluate the algorithms. One dataset included data for each site and growing season individually, and the other dataset was composed of data from sites and seasons.

The time interval of data used to train and evaluate the algorithms was from the beginning of the monitoring period, excluding the initial 24 h after installing the sensors (wet conditions) and data for GDD < 600 (plants not emerged yet or emerging), up to 2,200 GDD. No irrigation events were undertaken in the evaluated paddocks after 1,700 GDD. The selection of the parameter set used as algorithms inputs was based on two criteria: (i) the training response (lower errors), and (ii) one parameter cannot be a linear combination of the other input parameters. This last condition forbids having NDVI and ETc in the same input set. The selected input set for all evaluations was composed of GDD, ETc_n and $Rain_n$. The GDD index is relevant to locating the input data in the proper phenological crop stage. The other two parameters are essential to evaluate the soil water availability.

TABLE 3 Description of the machine learning models configuration.

Model	Layers/kernel	Optimizer
DMP	Input: 1 × 3 2 dense layers 32 neurons each – Tanh Output 1 × 1	Adam
SVR	Input: 1 × 3 Radial basis function (RBF) kernel Output 1 × 1	–

DMP, dense multilayer perceptron; SVR, support vector regression.

Machine learning algorithms

The machine learning models must learn the relation kPa/mm and forecast the soil matric potential in non-monitored areas giving one set of input parameters. As this relation is nonlinear, two models were selected to deal with this nonlinearity. First, the Support Vector Regression (SVR) algorithm is based on the Support Vector model to estimate the value of a point as the best hyperplane that represents the given sample (Smola and Schölkopf, 2004). The second algorithm was a Dense Multilayer Perceptron (DMP) neural network that can be understood as an ANN in which one neuron is connected with all neurons from the subsequent layers (Arnold et al., 2019). The SVR algorithm was implemented based on the Sklearn package, and the DMP model was developed using the Tensorflow API (Abadi et al., 2016).

Table 3 presents the configuration of each algorithm. All tests were performed using the k-fold method (Bengio and Grandvalet, 2004), which uses a cross-validation technique when the dataset is split in k folds with approximately the same number of samples. The algorithms are tested k times, each one changing the validation subset, i.e., each interaction deals with a selection of the $k - 1$ folds to training and one to test training, and each fold is used to test the algorithm once. The average of tested folds determines the algorithm's accuracy.

Before splitting data in folds, they were shuffled to mix the order of the points in the farm dataset and to mix data from different farms and seasons in the second type of dataset. In the last type of dataset, this is a way to avoid points from just one farm being part of a fold and data from other farms in the other fold. The risk of not doing this shuffle is to make the algorithms, particularly DMP, learn the relation kPa/mm related to one grower's practice instead of a pattern in the cotton crop.

The SVR and DMP models received the same training inputs format 1 × 3, i.e., one input has three parameters, and present the result as 1 × 1, i.e., one output with one estimated/forecasted point. The input set was a combination of parameters from Table 2, and the output was soil matric potential in all evaluations.

The R^2 (coefficient of determination) obtained for each tested fold was considered to evaluate the algorithms' training and responses. However, as the objective of the algorithm evaluation is not to estimate the exact soil matric potential but to provide an admissible value, the algorithms were also evaluated based on the percentage of estimates within an interval of 10 kPa. This interval was selected because of the fluctuation observed in soil matric potential between sensors installed in the same bay next to each other when soil matric potential data for all the sites was assessed.

Results

Relation between soil matric potential and ETc (kPa/mm)

The relation between soil matric potential and cumulative evapotranspiration between irrigations expressed in kPa/mm, represents the ratio of the soil matric potential and water demand according to the crop development. This relation can be expressed as a second-degree function (Figure 4) in which the inflexion point corresponds to the inflexion point from the soil matric potential chart (Figure 3). It allows the observation of the dynamics of soil matric potential even when soil is saturated, and the crop water demand cannot be readily evaluated in the soil matric potential chart. Figure 4 shows the relationship between soil matric potential and cumulative evapotranspiration in several irrigation events (one line for irrigation event). The R^2 of each relation kPa/mm is similar in farms A and C, even though irrigation management was different in each farm. In farm A irrigation scheduling was done based on the recommendations of an automatic irrigation control system while in farm A and B it was done based on the water managers' practices. The relation was not affected by weather events or the crop phenological stage. The same relation was observed at different phenological crop stages (see Figure 4 – farm B – irrigation event #4 in the end of the growing season when crop water demand is lower, and irrigation is not needed) and weather conditions. However, rainfall had a decreasing effect on the slope of such relation as can be observed in Figure 4 (farms A and B– irrigation event #4) when it remained flat for longer.

In farms A and B, the relation kPa/mm had a maximum value close to 2.5, while in farm C the maximum value was close to 4.5. The difference could be related to the soil and weather conditions that were different for farms A and B compared to farm C. Due to the difficulty of having an analytical model that encompasses such relation, the alternative is to apply a machine learning model. Such models may learn the relation kPa/mm and predict the soil matric potential through remote sensing data in areas not monitored by sensors.

Algorithms' evaluation in individual farms

The algorithms' evaluation in individual farms comprises the automatic irrigation in farm A and non-automatic (or traditional) irrigation strategies in farms B and C. Farms A and B present a smaller monitored area than farm C. Consequently, there are fewer points in farms A and B compared to farm C to train and evaluate resultant model from each algorithm. Due to this situation, farms A and B had only three folds to be evaluated, while farm C had five folds. Therefore, three folds from each farm are presented to evaluate the algorithms' responses in all following analyses.

The DMP presents R^2 above 0.80 in farms A and C (Figure 5), indicating the automation of the irrigation process did not play a decisive role in the algorithm estimation capability, which can be confirmed by SVR results (Figure 6). That SVR algorithm presents inferior performance compared to DMP in farms A and C, but the opposite in farm B, when the performance was substantially superior to DMP. It is possible to evaluate a sequence of points similar to a line in farm B (second and third folds in Figure 5) not closed from the R^2 line estimated by the DMP algorithm.

It is also evaluated the soil matric potential estimations/predictions versus measured points distribution in a ±10 kPa interval. Charts in Figures 7, 8 have the $R^2 = 1$ (black line) and the ±10 kPa represented between grey lines. Both charts represent individual farms with three folds each.

Both models' estimations below −50 kPa present equivalent results, while DMP presents better estimations for points below −50 kPa in farms A and C. The percentage of estimated points in the ±10 kPa interval; across all farms showed a satisfactory performance (Table 4). The DMP had better performance than SVR in farm C, while there is no difference between algorithms in farm A. The monitored area in farm C is more extensive than on farms A and B. Due to the size of the bays in farm C, an irregular moisture distribution in the bays could be observed on the same day, which means the algorithms had to deal with significant differences in soil matric potential corresponding to the same GDD values in the input parameters.

Farm A presents the same results for both algorithms reflecting the regularity promoted by the automation process in the irrigated areas. However, farm B presented a significant difference between algorithms with SVR superior to DMP, getting 90% of corrected estimated points in the ±10 kPa band (see farm B – SVR dispersion in Figure 8).

Evaluation of algorithms considering all farms and growing seasons combined

In individual farm evaluation, the differences in irrigation strategies did not cause overfitting in the neural network model but could cause a bias in the results (grower practices). Combining the measured points from all farms and seasons and shuffling the

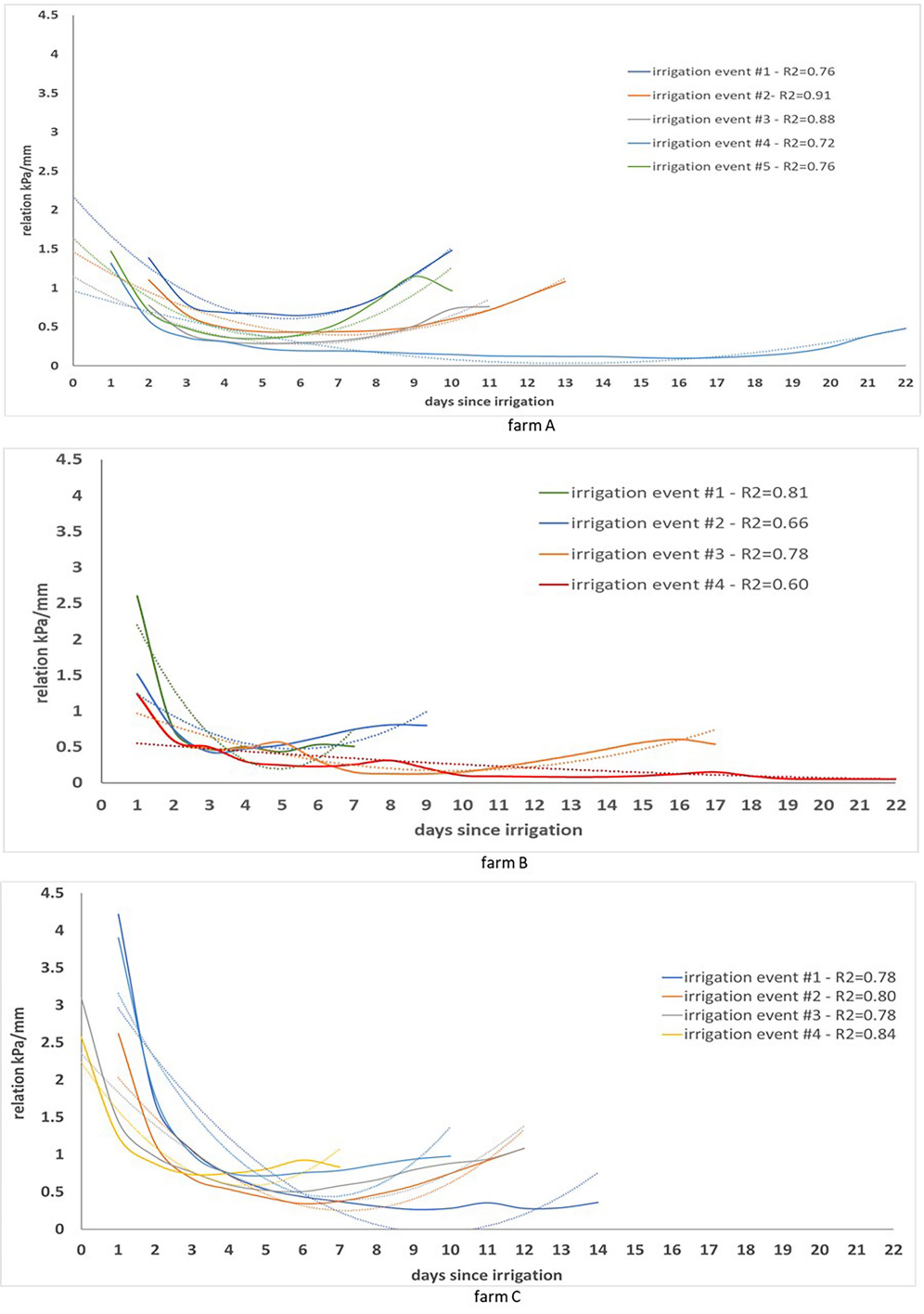


FIGURE 4
Evolution of the relation kPa/mm observed during several irrigation events at each farm. Within each farm, the least square function and coefficient of determination (R^2) is shown for each irrigation event.

points are essential to creating the evaluation dataset when such bias is not present.

Figure 9 presents the soil matric potential measured versus estimated to evaluate the responses provided by SVR and DMP. Both algorithms had equivalent results in all testing folds. According to the R^2 metric, both algorithms present similar learning capabilities. The DMP reached $R^2 = 0.8424$ in the fold (d), and SVR reached $R^2 = 0.7559$ in the fold (a), meaning the DMP model provided a more accurate estimation of the soil matric potential in the root zone with a given input set.

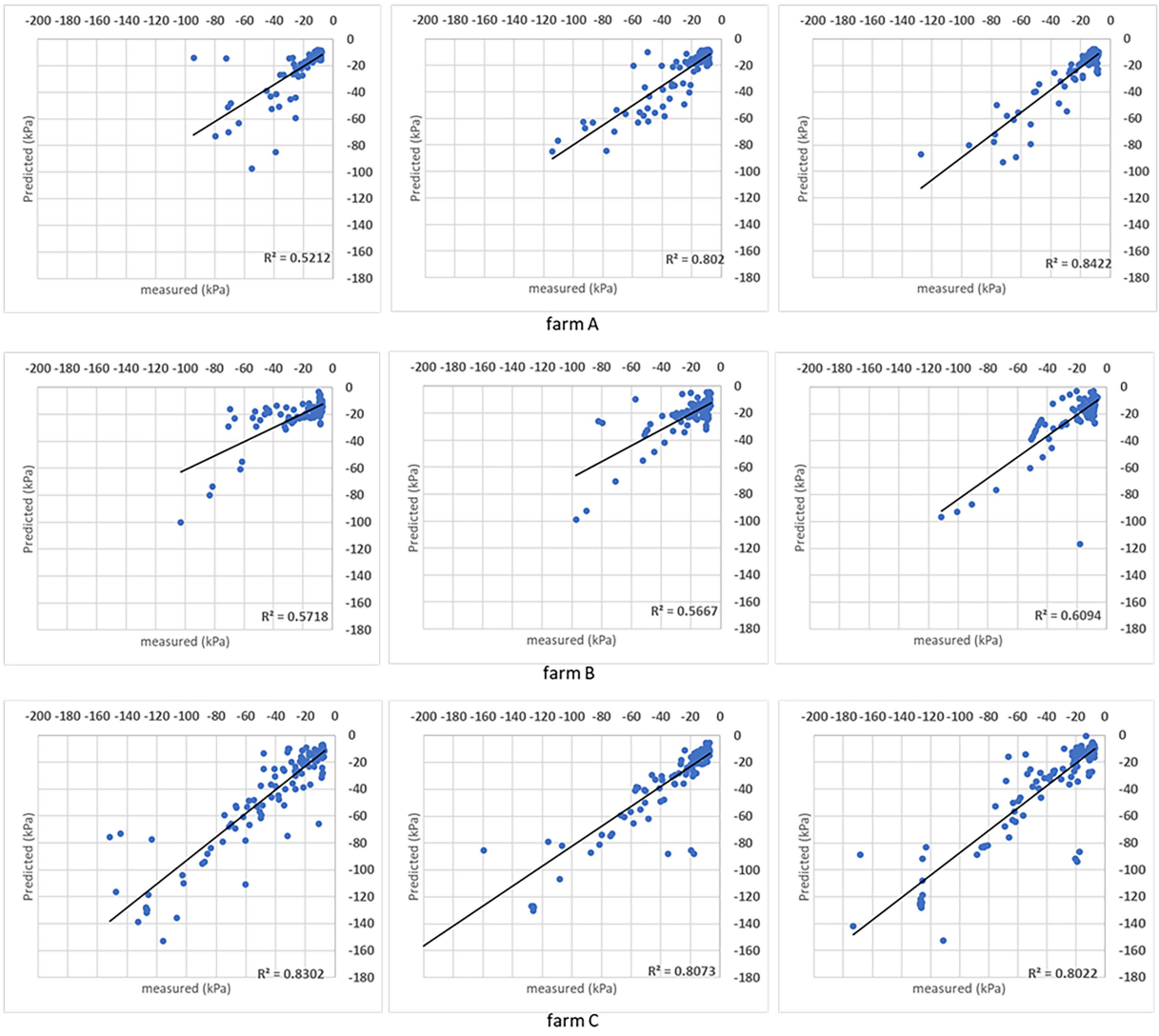


FIGURE 5
Results obtained with the DMP model when the analysis was done for individual sites. For each farm, the coefficient of determination (R^2) obtained for the comparison between measured and predicted soil matric potential is shown for three folds.

The performance of each model considering the ±10 kPa interval is represented in Figure 10. Data organization followed the k-fold method that generated other sets of folds for training/evaluation sessions. With the reorganization of points, the same characteristics were found and represented in Figure 9 (compared to Figure 8). Both models present similar responses when estimated points are contained in the proposed interval, reflecting the field's measurements. Evaluating how many estimated points are in the ±10 kPa (Table 4), the SVR presents slightly better results than DMP. On average, 69.56% of points in DMP and 72.98% of points in SVR are in the proposed interval. Increasing the interval to ±15 kPa, both models present the same results (~80%) since the interval includes more correctly predicted points below −60 kPa.

Discussion

Relationship between soil matric potential and ET_{cn}

Soil water dynamics are influenced by many bio-physical factors forming a complex system that is difficult to capture by analytical models without a significant number of input parameters and a complete understanding of their relationship. However, changes in soil moisture storage can be determined by a soil water balance model as the difference between water added to the soil (precipitation and irrigation) and water lost in the system (deep percolation, run off, lateral flow and evapotranspiration; O'Geen, 2013). In the context of this study, decreases in soil moisture storage in the top 0.20 m of soil are mainly due to the

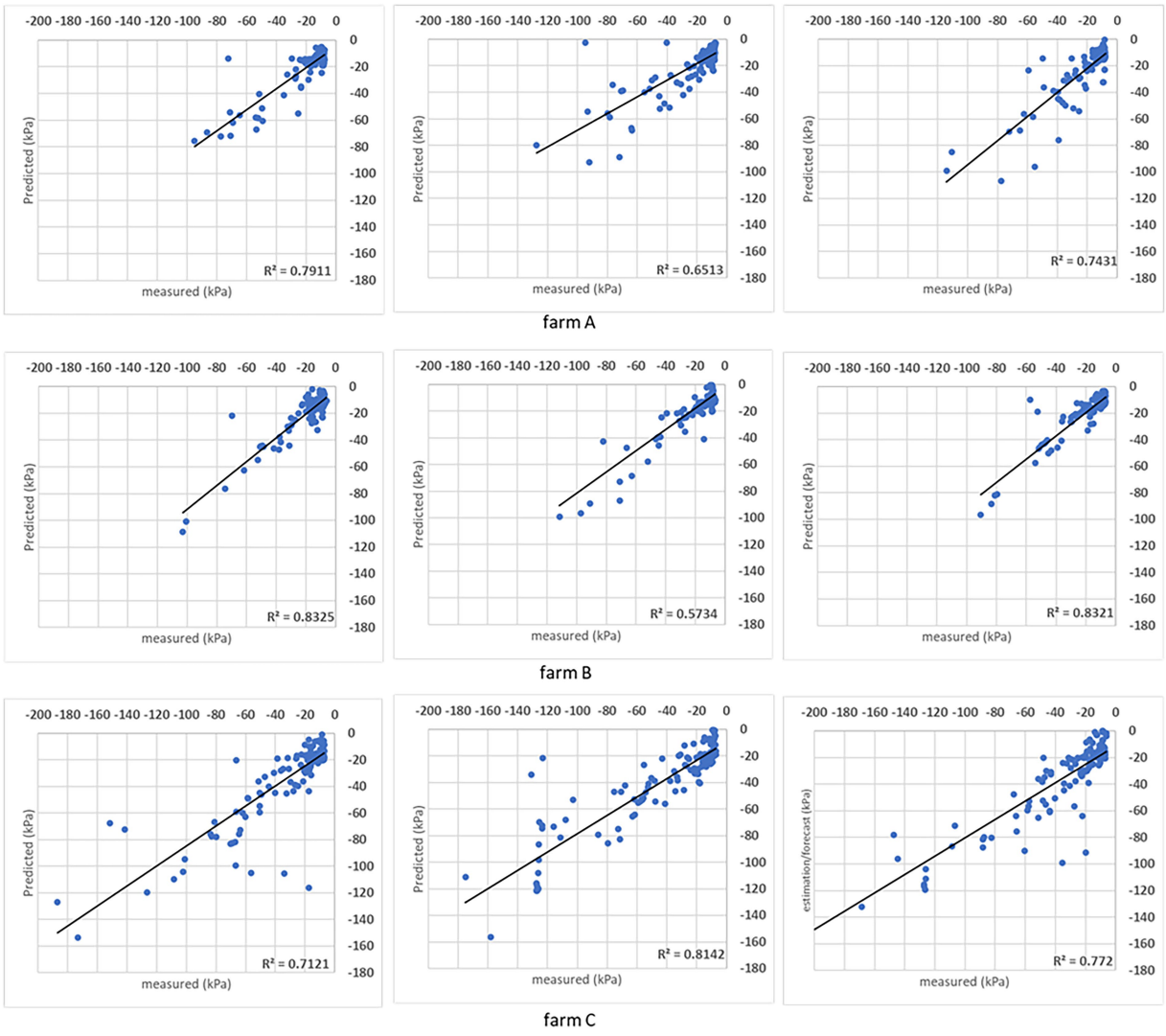


FIGURE 6
Results obtained with the SVR model when the analysis was done for individual sites. For each farm, the coefficient of determination (R^2) obtained for the comparison between measured and predicted soil matric potential is shown for three folds.

crop evapotranspiration. The relationship between soil matric potential and cumulative crop evapotranspiration between irrigation events (kPa/mm) that is expressed by a second-degree function should then represent this system. The R^2 obtained for this relationship was in most cases >0.72 even though data was collected at different times in the season and from farms with different soil characteristics, water management strategies and weather conditions (Figure 4). For example, in farms A and B, the 2019/20 cotton growing season was hotter and dryer than the 2020/21 growing season. The highest (47°C) and lowest (19°C) maximum temperatures for the month of January since 1960 were recorded in the 2019/20 and 2020/21 growing seasons, respectively. These results suggest that the soil matric potential-ET_{cn} relation between irrigation events is stable and machine learning models that capture this relationship can potentially be used to estimate soil matric potential from ET_c measurements.

Performance of the models for individual and combined datasets

The SVR and DMP models were used in this study to learn the pattern of the proposed relationship to be able to indirectly estimate soil matric potential in areas not monitored with sensors from satellite-NDVI derived ET_c measurements. This was possible without using specific soil characteristics as data input as it was required in other studies such as Villani et al. (2018). The approach followed in this work is different from other studies using remote sensed data as an input parameter in machine learning models for the estimation or prediction of soil moisture (Ahmad et al., 2010; Coopersmith et al., 2016; Torres-Rua et al., 2016; Sungmin and Rene, 2021). In Ahmad et al. (2010), microwave backscatter observations and incidence angle from Tropical Rainfall Measuring Mission (TRMM), NDVI and

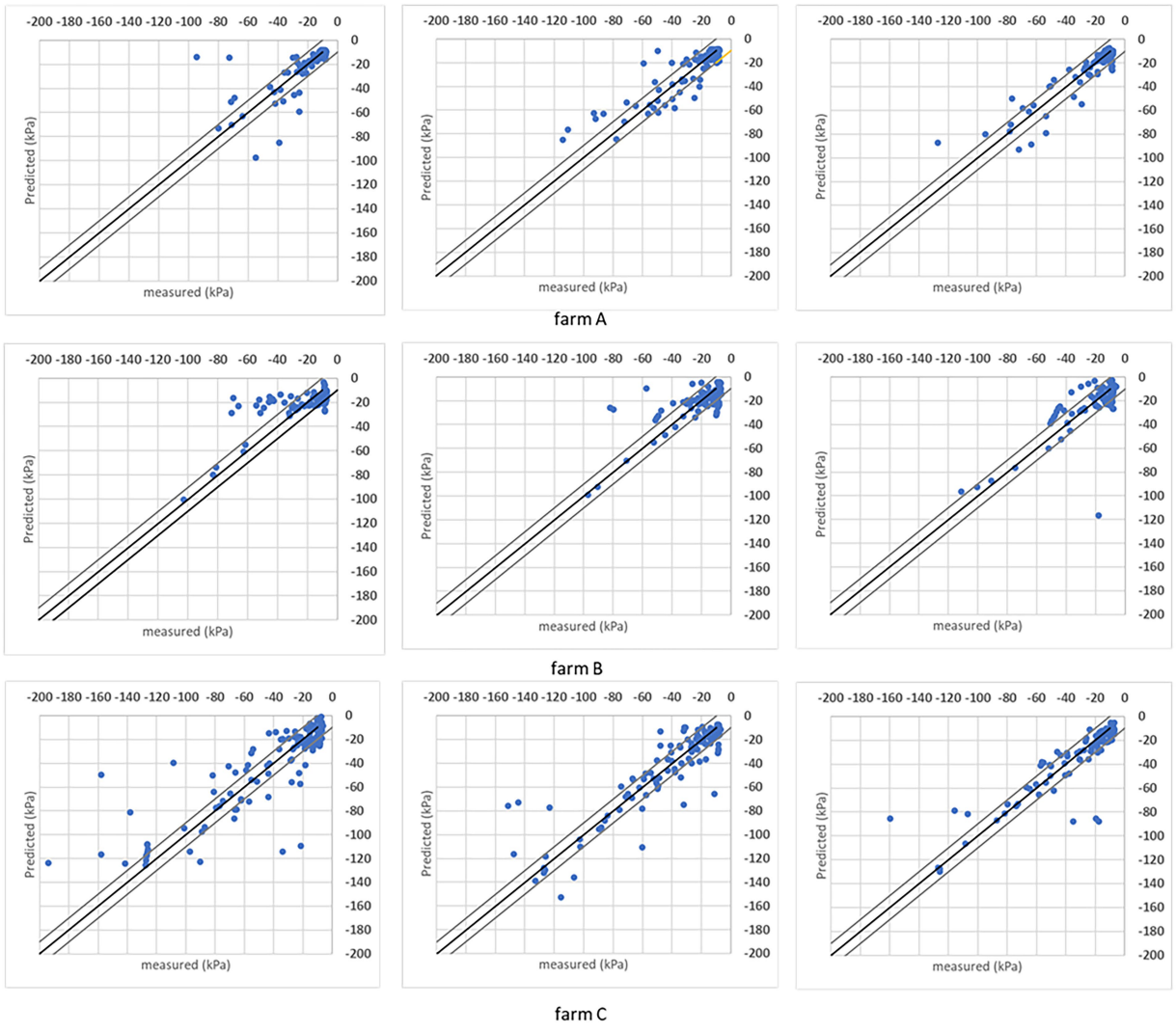


FIGURE 7
Results obtained with the DMP model when the analysis was done for individual sites considering the R^2=1 (black line) and±10kPa interval between the grey lines. Three folds are shown for each farm.

simulated soil moisture data (at 0.10 m depth) were used as input parameters in a Support Vector Machine model to estimate soil moisture with good results. In Torres-Rua et al. (2016), surface soil moisture was estimated at spatial and temporal resolution using a Relevance Vector Machine model by combining *in situ* soil moisture and weather data with satellite-derived evapotranspiration (METRIC model). These models used in Ahmad et al. (2010) and Torres-Rua et al. (2016), were effective in estimating volumetric water content in the top 0.05 m of soil but its feasibility for estimating soil moisture at deeper soil layers was not studied which limits the applicability of these models to irrigation decision support systems. In the study here presented, soil matric potential was preferred to soil water content because of its suitability for irrigation automation and unneeded soil-specific calibration. The SVR and DMP models learnt the pattern of the soil matric potential and crop evapotranspiration relationship using the GDD index, satellite NDVI-derived ET_c and rainfall as input parameters. The models' output was the estimated soil matric potential at 0.20 m depth, where thresholds can be used in practice to trigger irrigation events. Gumiere et al. (2020) and Dubois et al. (2021) also explored machine learning models for predicting soil matric potential in cranberry and potato crops, respectively, for irrigation management. In Gumiere et al. (2020), a Random Forest (RF) model with rainfall, reference evapotranspiration and soil matric potential measurements at 0.10 m as input data predicted hourly soil matric potential with an R^2 of 0.58. The average R^2 obtained with the DMP and SVR models for individuals farms and for the entire dataset combined was ≥0.70 in this study. The best performance of the models here tested could be related with the fact that Gumiere et al. (2020) predicted soil matric potential hourly while daily values were estimated in this work. The RF, SVR, and Neural Network (NN)

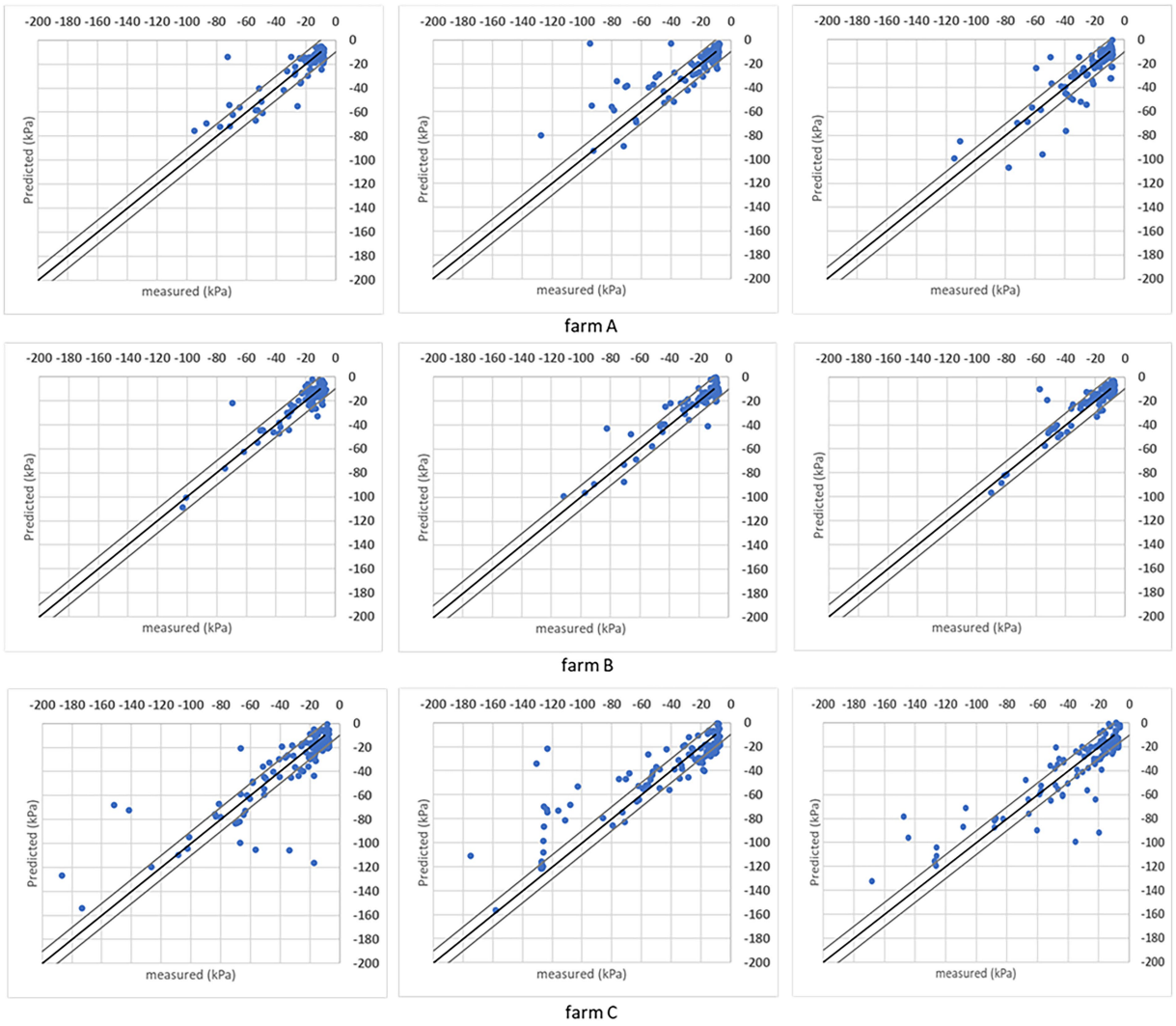


FIGURE 8
Results obtained with the SVR model when the analysis was done for individual sites considering the R^2=1 (black line) and±10kPa interval between the grey lines. Three folds are shown for each farm.

models used in Dubois et al. (2021) had a higher performance (R^2 > 0.92) in predicting soil matric potential than the models used here although remotely sensed data was not used as an input parameter in their study and thus it cannot be used for the purpose proposed in this study.

Soil moisture estimation by machine learning models

Both SVR and DMP models performed well in estimating soil matric potential (Tables 4, 5). This is a strong indication that the kPa/mm relation can be used to estimate the soil matric potential in the root zone of non-monitored areas from NDVI data, rainfall, and the GDD index that represents the development stage of the crop. The performance assessment for individual datasets (farms) showed that both models performed similarly and that they performed slightly better in farm A and B than in farm C. In particular, the SVR

TABLE 4 Percentage of estimated points between the ±10kPa and ±15kPa intervals in the analysis for all farms and seasons combined applying the DMP and SVR models.

Fold#	Points in ±10 kPa band		Points in ±15 kPa band	
	DMP	SVR	DMP	SVR
1	0.67	0.77	0.76	0.83
2	0.65	0.68	0.78	0.76
3	0.71	0.72	0.81	0.82
4	0.75	0.74	0.83	0.83
5	0.69	0.74	0.82	0.83
Avg.	0.69	0.73	0.80	0.82

DMP, dense multilayer perceptron; SVR, support vector regression.

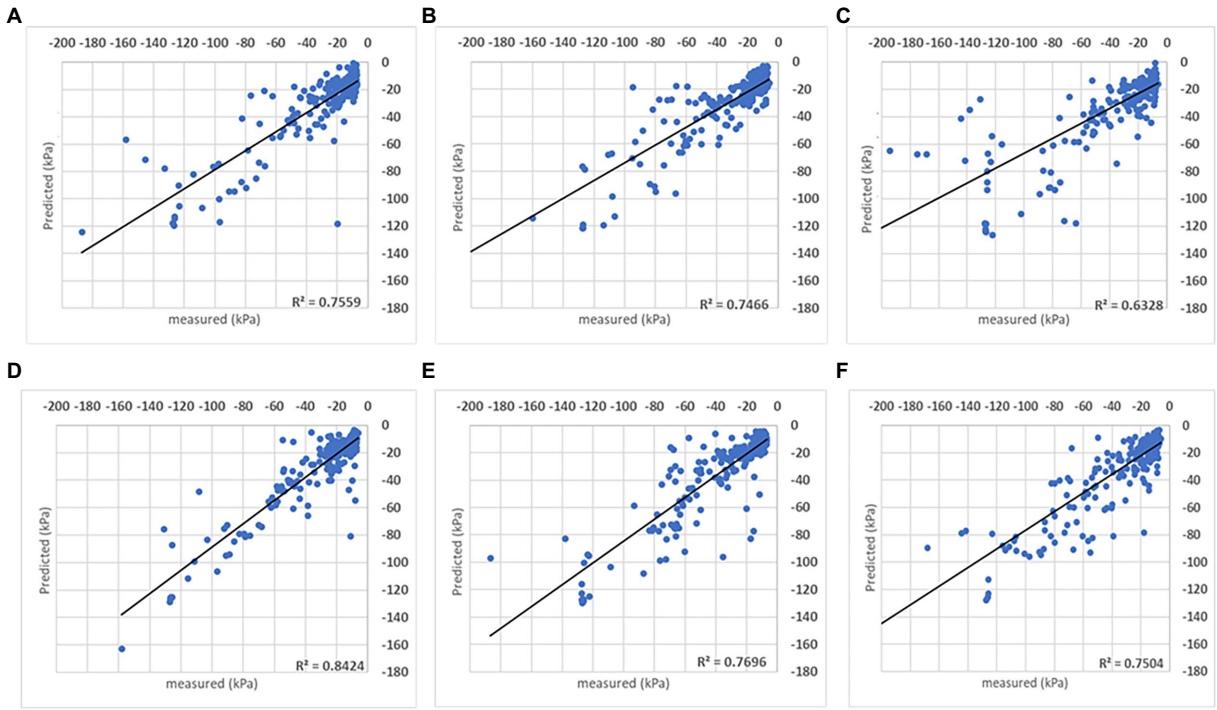


FIGURE 9
Results obtained with the SVR **(A–C)** and DMP **(D–F)** algorithms when the analysis was done combining data from all sites and seasons. The coefficient of determination (R^2) obtained for the comparison between measured and predicted soil matric potential is shown for three folds.

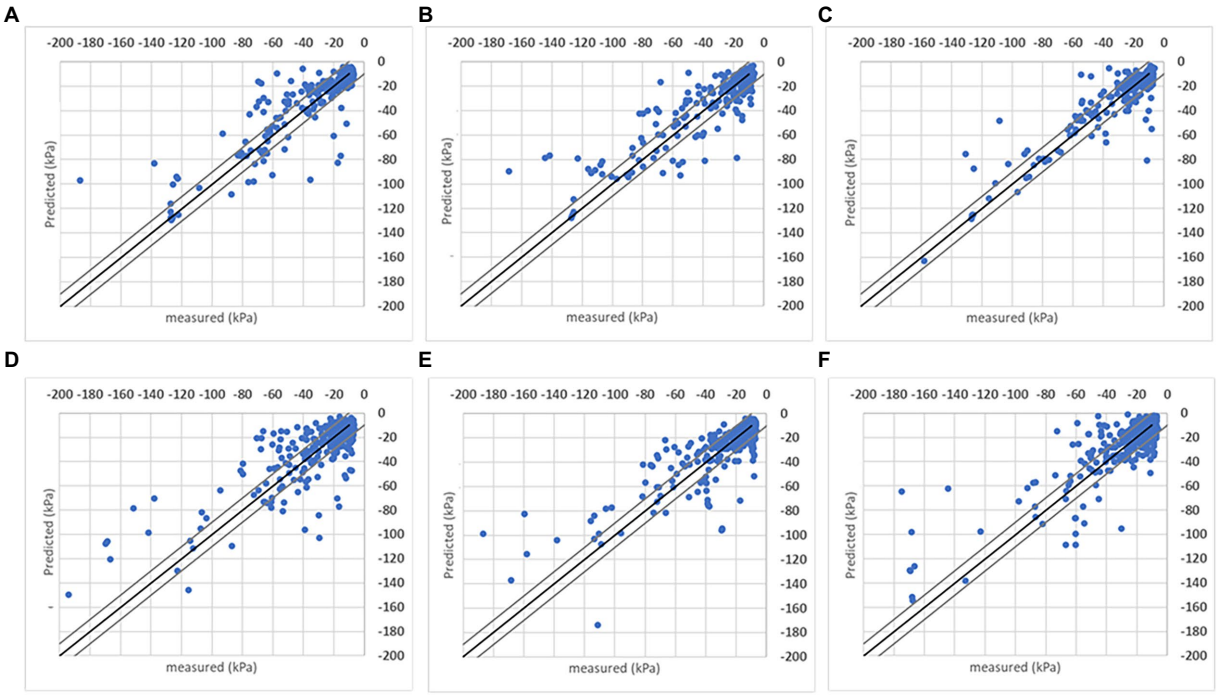


FIGURE 10
Results obtained with the SVR **(A–C)** and DMP **(D–F)** algorithms when the analysis was done combining data from all sites and seasons considering the R^2=1 (black line) and±10kPa interval between the grey lines. Three folds are shown for each farm.

TABLE 5 Percentage of estimated points between the ±10kPa interval in the analysis for individual farms with the DMP and SVR models.

Fold#	Farm A		Farm B		Farm C	
	DMP	SVR	DMP	SVR	DMP	SVR
1	0.89	0.88	0.75	0.94	0.55	0.63
2	0.78	0.80	0.77	0.87	0.60	0.61
3	0.83	0.82	0.78	0.90	0.78	0.64
4	–	–	–	–	0.80	0.57
5	–	–	–	–	0.84	0.58
Avg.	0.84	0.84	0.77	0.90	0.72	0.61

DMP, dense multilayer perceptron; SVR, support vector regression.

model in farm B presented the highest accuracy, but this could just be that the models fit better the dataset for those particular farm and season. The performance of the models considering the ±10 kPa interval when data from all sites and seasons was combined was slightly worse than for individual farms although still with an accuracy around 70% (Table 4). The SVR model estimated 4% more readings within the ±10 kPa interval than DMP. The performance improved significantly in both models when the interval increased to ±15 kPa mainly due to the inclusion of estimates of soil matric potential between −60 kPa and −120 kPa within this larger interval. The dataset used for the training and testing of the models had more readings between 0 and −60 kPa than below this value because irrigation was scheduled in these sites to ensure plant water availability and avoid water stress conditions. Consequently, the soil matric potential dataset below −60 kPa was scarce for a proper training of the models and the accuracy in estimating soil matric potential in drier soil was lower.

Conclusion

This study showed that soil matric potential and cumulative ET_c between irrigation events (kPa/mm) have a stable and robust relationship that integrates the effects of soil type and weather condition in a second-degree function. It also demonstrated that machine learning models with capability to process time-dependent parameters such as the DMP and SVR applied here can learn the patter of the soil matric potential-ET_{cn} relation. This offers the possibility of accurately estimating soil matric potential in the root zone of crops in non-monitored areas with in-filed sensors from remotely sensed ET_c estimates. The approach is scalable to farms with multiple irrigation fields without the limitations of on-ground sensing related to the cost and organization of the sensors network.

The assessment of this relation kPa/mm with machine learning models provides a new technique to estimate soil tension in the root zone of cotton crops although it is potentially suitable for other crops. Future work can consider collecting data across wider areas and more seasons to refine the relation kPa/mm and machine learning models' performance. Additionally, research on expanding the relation kPa/mm in other crops to evaluate if the relation sustains and how well it represents the complexity of the crop/soil/water/weather system has merit for other industries.

Data availability statement

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

Author contributions

JH, CL, and RM: project conceptualization, data evaluation and validation, and review and editing. RM and CL: methodology and writing. RM: relation and machine learning models. JH: funding acquisition. All authors contributed to the article and approved the submitted version.

Funding

This work is supported by funding from the Australian Government Department of Agriculture, Water and the Environment as part of its R&D for Profit Program and the Cotton Research and Development Corporation.

Acknowledgments

We would like to thank Brenno Tondato de Faria for the machine learning algorithms coding suggestions and library discussions during the development of the software used in this work.

Conflict of interest

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

Publisher's note

References

Abadi, M., Barham, P., Chen, J., Chen, Z., Davis, A., Dean, J., et al. (2016). "{TensorFlow}: a system for {Large-Scale} machine learning", in: 12th USENIX symposium on operating systems design and implementation (OSDI 16), 265–283.

Ahmad, S., Kalra, A., and Stephen, H. (2010). Estimating soil moisture using remote sensing data: a machine learning approach. *Adv. Water Resour.* 33, 69–80. doi: 10.1016/j.advwatres.2009.10.008

Akshay, S., and Ramesh, T. (2020). "Efficient machine learning algorithm for smart irrigation", in: 2020 International Conference on Communication and Signal Processing (ICCSP): IEEE, 867–870.

Allen, R. G., Pereira, L. S., Raes, D., and Smith, M. (1998). Crop evapotranspiration-guidelines for computing crop water requirements-FAO irrigation and drainage paper 56. FAO, Rome 300, D05109.

Alzubi, J., Nayyar, A., and Kumar, A. (2018). Machine learning from theory to algorithms: an overview. *J. Phys. Conf. Ser.* 1142:012012. doi: 10.1088/1742-6596/1142/1/012012

Arnold, T., Kane, M., and Lewis, B. W. (2019). *A Computational Approach to Statistical Learning*. Milton, UK: CRC Press LLC, doi: 10.1201/9781315171401

Babaeian, E., Paheding, S., Siddique, N., Devabhaktuni, V. K., and Tuller, M. (2021). Estimation of root zone soil moisture from ground and remotely sensed soil information with multisensor data fusion and automated machine learning. *Remote Sens. Environ.* 260:112434. doi: 10.1016/j.rse.2021.112434

Ballester, C., Hornbuckle, J., Brinkhoff, J., and Quayle, W. C. (2021). Effects of three frequencies of irrigation and nitrogen rates on lint yield, nitrogen use efficiency and fibre quality of cotton under furrow irrigation. *Agric. Water Manag.* 248:106783. doi: 10.1016/j.agwat.2021.106783

Bengio, Y., and Grandvalet, Y. (2004). No unbiased estimator of the variance of k-fold cross-validation. *J. Mach. Learn. Res.* 5, 1089–1105.

Benos, L., Tagarakis, A. C., Dolias, G., Berruto, R., Kateris, D., and Bochtis, D. (2021). Machine learning in agriculture: a comprehensive updated review. *Sensors* 21:3758. doi: 10.3390/s21113758

Bhoi, A., Nayak, R. P., Bhoi, S. K., and Sethi, S. (2021a). "*Automated Precision Irrigation System Using Machine Learning and IoT*". (Singapore: Springer), 275–282.

Bhoi, A., Nayak, R. P., Bhoi, S. K., Sethi, S., Panda, S. K., Sahoo, K. S., et al. (2021b). IoT-IIRS: internet of things based intelligent-irrigation recommendation system using machine learning approach for efficient water usage. *PeerJ Computer Science* 7:e578. doi: 10.7717/peerj-cs.578

Bittelli, M. (2011). Measuring soil water content: a review. *HortTechnology* 21, 293–300. doi: 10.21273/HORTTECH.21.3.293

Blasch, J., Van Der Kroon, B., Van Beukering, P., Munster, R., Fabiani, S., Nino, P., et al. (2022). Farmer preferences for adopting precision farming technologies: a case study from Italy. *Eur. Rev. Agric. Econ.* 49, 33–81. doi: 10.1093/erae/jbaa031

Brinkhoff, J., Hornbuckle, J., Quayle, W., Lurbe, C. B., and Dowling, T. (2017). "WiField, an IEEE 802.11-based agricultural sensor data gathering and logging platform", in: 2017 Eleventh International Conference on Sensing Technology (ICST): IEEE), 1–6.

Cahn, M. D., and Johnson, L. F. (2017). New approaches to irrigation scheduling of vegetables. *Horticulturae* 3:28. doi: 10.3390/horticulturae3020028

Campoverde, L. M. S., Tropea, M., and De Rango, F. (2021). "An IoT based smart irrigation management system using reinforcement learning modeled through a markov decision process", in: 2021 IEEE/ACM 25th International Symposium on Distributed Simulation and Real Time Applications (DS-RT): IEEE, 1–4.

Cardoso, J., Glória, A., and Sebastiao, P. (2020). "Improve irrigation timing decision for agriculture using real time data and machine learning", in: 2020 International Conference on Data Analytics for Business and Industry: Way Towards a Sustainable Economy (ICDABI): IEEE, 1–5.

Chaudhary, S. K., and Srivastava, P. K. (2021). Future challenges in agricultural water management. *Agri. Water Manag.* 445–456. doi: 10.1016/B978-0-12-812362-1.00021-7

Coopersmith, E. J., Cosh, M. H., Bell, J. E., and Boyles, R. (2016). Using machine learning to produce near surface soil moisture estimates from deeper in situ records at US climate reference network (USCRN) locations: analysis and applications to AMSR-E satellite validation. *Adv. Water Resour.* 98, 122–131. doi: 10.1016/j.advwatres.2016.10.007

Cosgrove, W. J., and Loucks, D. P. (2015). Water management: current and future challenges and research directions. *Water Resour. Res.* 51, 4823–4839. doi: 10.1002/2014WR016869

Dubois, A., Teytaud, F., and Verel, S. (2021). Short term soil moisture forecasts for potato crop farming: a machine learning approach. *Comput. Electron. Agric.* 180:105902. doi: 10.1016/j.compag.2020.105902

FAO (2017). *The Future of Food and Agriculture—Trends and Challenges*. Rome: FAO.

Filev Maia, R., Ballester Lurbe, C., Agrahari Baniya, A., and Hornbuckle, J. (2020). IRRISENS: an IoT platform based on microservices applied in commercial-scale crops working in a multi-cloud environment. *Sensors* 20:7163. doi: 10.3390/s20247163

González-Briones, A., Castellanos-Garzón, J. A., Mezquita-Martín, Y., Prieto, J., and Corchado, J. M. (2019). "A multi-agent system framework for autonomous crop irrigation", in: 2019 2nd International Conference on Computer Applications & Information Security (ICCAIS): IEEE, 1–6.

Grabham, M. (2012). "Bankless channel irrigation systems," in *WATERpack—A Guide for Irrigation Management in Cotton and Grain Farming Systems*. (*3rd Edn.*). ed. DW Consulting Services (Narrabri, Australia: Cotton Research and Development Corporation (CRDC)), 388–391.

Greifeneder, F., Notarnicola, C., and Wagner, W. (2021). A machine learning-based approach for surface soil moisture estimations with google earth engine. *Remote Sens.* 13:2099. doi: 10.3390/rs13112099

Grillakis, M. G., Koutroulis, A. G., Alexakis, D. D., Polykretis, C., and Daliakopoulos, I. N. (2021). Regionalizing root-zone soil moisture estimates from ESA CCI soil water index using machine learning and information on soil, vegetation, and climate. *Water Resour. Res.* 57:e2020WR029249. doi: 10.1029/2020WR029249

Gumiere, S. J., Camporese, M., Botto, A., Lafond, J. A., Paniconi, C., Gallichand, J., et al. (2020). Machine learning vs. physics-based modeling for real-time irrigation management. *Front. Water* 2:8. doi: 10.3389/frwa.2020.00008

Ikidid, A., Fazziki, A. E., and Sadgal, M. (2021). "Smart collective irrigation: agent and internet of things based system", in: Proceedings of the 13th International Conference on Management of Digital EcoSystems, 100–106.

Irrometer (2021). Reading WATERMARK soil moisture sensors [online]. Available at: https://www.irrometer.com/200ss.html [Accessed January 2, 2022].

Jägermeyr, J. (2020). Agriculture's historic twin-challenge toward sustainable water use and food supply for all. *Front. Sustain. Food Sys.* 4:35. doi: 10.3389/fsufs.2020.00035

Jones, H. G. (2006). Monitoring plant and soil water status: established and novel methods revisited and their relevance to studies of drought tolerance. *J. Exp. Bot.* 58, 119–130. doi: 10.1093/jxb/erl118

Kang, Y., Wang, R., Wan, S., Hu, W., Jiang, S., and Liu, S. (2012). Effects of different water levels on cotton growth and water use through drip irrigation in an arid region with saline ground water of Northwest China. *Agric. Water Manag.* 109, 117–126. doi: 10.1016/j.agwat.2012.02.013

Kondaveti, R., Reddy, A., and Palabtla, S. (2019). "Smart irrigation system using machine learning and IOT", in: 2019 International Conference on Vision Towards Emerging Trends in Communication and Networking (ViTECoN): IEEE, 1–11.

Liakos, K. G., Busato, P., Moshou, D., Pearson, S., and Bochtis, D. (2018). Machine learning in agriculture: a review. *Sensors* 18:2674. doi: 10.3390/s18082674

McMahon, J., and Low, A. (1972). Growing degree days as a measure of temperature effects on cotton. *Cotton Growing Rev.* 49, 39–49.

Murthy, A., Green, C., Stoleru, R., Bhunia, S., Swanson, C., and Chaspari, T. (2019). "Machine learning-based irrigation control optimization", in: Proceedings of the 6th ACM International Conference on Systems for Energy-Efficient Buildings, Cities, and Transportation, 213–222.

O'Geen, A. T. (2013). *Soil Water Dynamics Nature Education Knowledge*. 4:9. Available at: https://www.nature.com/scitable/knowledge/library/soil-water-dynamics-103089121/

Orth, R. (2021). Global soil moisture data derived through machine learning trained with in-situ measurements. *Sci. Data* 8, 1–14.

Paloscia, S., Pettinato, S., Santi, E., Notarnicola, C., Pasolli, L., and Reppucci, A. (2013). Soil moisture mapping using Sentinel-1 images: algorithm and preliminary validation. *Remote Sens. Environ.* 134, 234–248. doi: 10.1016/j.rse.2013.02.027

Perea, R. G., Ballesteros, R., Ortega, J. F., and Moreno, M. Á. (2021). Water and energy demand forecasting in large-scale water distribution networks for irrigation using open data and machine learning algorithms. *Comput. Electron. Agric.* 188:106327. doi: 10.1016/j.compag.2021.106327

Pereira, L. S., Allen, R. G., Smith, M., and Raes, D. (2015). Crop evapotranspiration estimation with FAO56: past and future. *Agric. Water Manag.* 147, 4–20. doi: 10.1016/j.agwat.2014.07.031

Pimentel, D., Berger, B., Filiberto, D., Newton, M., Wolfe, B., Karabinakis, E., et al. (2004). Water resources: agricultural and environmental issues. *Bioscience* 54, 909–918. doi: 10.1641/0006-3568(2004)054[0909:WRAAEI]2.0.CO;2

Pôças, I., Calera, A., Campos, I., and Cunha, M. (2020). Remote sensing for estimating and mapping single and basal crop coefficients: a review on spectral

vegetation indices approaches. *Agric. Water Manag.* 233:106081. doi: 10.1016/j.agwat.2020.106081

Ponraj, A. S., and Vigneswaran, T. (2020). Daily evapotranspiration prediction using gradient boost regression model for irrigation planning. *J. Supercomput.* 76, 5732–5744. doi: 10.1007/s11227-019-02965-9

Prasad, R., Deo, R. C., Li, Y., and Maraseni, T. (2018). Soil moisture forecasting by a hybrid machine learning technique: ELM integrated with ensemble empirical mode decomposition. *Geoderma* 330, 136–161. doi: 10.1016/j.geoderma.2018.05.035

Saad, A., and Gamatié, A. (2020). Water management in agriculture: a survey on current challenges and technological solutions. *IEEE Access* 8, 38082–38097. doi: 10.1109/ACCESS.2020.2974977

Shunlin Liang, J. W. (2020). "Chapter 18. Soil moisture contents," in *Advanced Remote Sensing. 2nd Edn*. S. Liang and J. Wang (New York: Academic Press), 685–711.

Singh, G., Sharma, D., Goap, A., Sehgal, S., Shukla, A., and Kumar, S. (2019). "Machine learning based soil moisture prediction for internet of things based smart irrigation system", in: 2019 5th International Conference on Signal Processing, Computing and Control (ISPCC): IEEE, 175–180.

Smola, A. J., and Schölkopf, B. (2004). A tutorial on support vector regression. *Stat. Comput.* 14, 199–222. doi: 10.1023/B:STCO.0000035301.49549.88

Sungmin, O., and Rene, O. (2021). Global soil moisture data derived through machine learning trained with in-situ measurements. *Sci. Data* 8:170. doi: 10.1038/s41597-021-00964-1

Susha Lekshmi, S. U., Singh, D. N., and Shojaei Baghini, M. (2014). A critical review of soil moisture measurement. *Measurement* 54, 92–105. doi: 10.1016/j.measurement.2014.04.007

Thompson, R., Gallardo, M., Valdez, L., and Fernández, M. (2007). Using plant water status to define threshold values for irrigation management of vegetable crops using soil moisture sensors. *Agric. Water Manag.* 88, 147–158. doi: 10.1016/j.agwat.2006.10.007

Torres-Rua, A. F., Ticlavilca, A. M., Bachour, R., and Mckee, M. (2016). Estimation of surface soil moisture in irrigated lands by assimilation of landsat vegetation indices, surface energy balance products, and relevance vector machines. *Water* 8:167. doi: 10.3390/w8040167

Trout, T. J., and Johnson, L. F. (2007). "Estimating crop water use from remotely sensed NDVI, crop models, and reference ET", in: USCID Fourth International Conference on Irrigation and Drainage, Sacramento, California, 275–285.

Villani, G., Castaldi, P., Toscano, A., Stanghellini, C., Cinotti, T. S., Maia, R. F., et al. (2018). "Soil water balance model CRITERIA-ID in SWAMP project: proof of concept", in: 2018 23rd Conference of Open Innovations Association (FRUCT), 398–404.

frontiers | Frontiers in Plant Science

TYPE Methods
PUBLISHED 07 November 2022
DOI 10.3389/fpls.2022.897883

OPEN ACCESS

EDITED BY
Lei Shu,
Nanjing Agricultural University, China

REVIEWED BY
Jun Meng,
Dalian University of Technology, China
Lifu Gao,
Hefei Institutes of Physical Science (CAS), China
Chuanlei Zhang,
Tianjin University of Science and Technology, China

*CORRESPONDENCE
Shaowen Li
shwli@ahau.edu.cn

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

RECEIVED 16 March 2022
ACCEPTED 11 October 2022
PUBLISHED 07 November 2022

CITATION
Fu L, Li S, Rao Y, Liang J, Teng J and He Q (2022) A novel heuristic target-dependent neural architecture search method with small samples. *Front. Plant Sci.* 13:897883. doi: 10.3389/fpls.2022.897883

A novel heuristic target-dependent neural architecture search method with small samples

Leiyang Fu[1,2], Shaowen Li[1,2]*, Yuan Rao[1,2], Jinxin Liang[1,2], Jie Teng[1,2] and Quanling He[1,2]

[1]School of Information and Computer Science, Anhui Agricultural University, Hefei, Anhui, China, [2]Department of Science and Technology, Anhui Provincial Key Laboratory of Smart Agricultural Technology and Equipment, Hefei, Anhui, China

It is well known that crop classification is essential for genetic resources and phenotype development. Compared with traditional methods, convolutional neural networks can be utilized to identify features automatically. Nevertheless, crops and scenarios are quite complex, which makes it challenging to develop a universal classification method. Furthermore, manual design demands professional knowledge and is time-consuming and labor-intensive. In contrast, auto-search can create network architectures when faced with new species. Using rapeseed images for experiments, we collected eight types to build datasets (rapeseed dataset (RSDS)). In addition, we proposed a novel target-dependent search method based on VGGNet (target-dependent neural architecture search (TD-NAS)). The result shows that test accuracy does not differ significantly between small and large samples. Therefore, the influence of the dataset size on generalization is limited. Moreover, we used two additional open datasets (Pl@ntNet and ICL-Leaf) to test and prove the effectiveness of our method due to three notable features: (a) small sample sizes, (b) stable generalization, and (c) free of unpromising detections.

KEYWORDS

crop classification, target-dependent, neural architecture search, small samples, Bayesian optimization

Introduction

Image classification can distinguish objects by color, texture, shape, and spatial relationship. It uses computers to analyze images and classify each pixel or region into several categories without human interpretation (Wang and Wang, 2019). The following are two agricultural scenarios. (a) Genetic resources: artificial recognition is time-consuming and near-impossible. Automatic species identification is significant for

taxonomy. Purohit et al. (2016) studied a machine-learning method using leaf characteristics to recognize species. In recent years, deep learning, which can automatically extract features from original data, has dramatically improved classification performance (Barré et al., 2017; Lee and Chang, 2017; Pawara et al., 2017). Some studies propose associating machine learning with neural networks (Fu et al., 2016; Li et al., 2020). (b) Phenotype development: a phenotype is a characteristic or combination of an organism influenced by the genotype and by the environment. Usually, plants grow in a highly variable environment. More accurate and robust algorithms are needed to deal with complex backgrounds and quantify phenotypic characteristics. They can distinguish different components or even instances. Li et al. (2020) thoroughly reviewed phenotyping technologies and used machine vision to measure plant stress.

There exist several machine-learning methods such as support vector machines (SVMs) (Rumpf et al., 2010), k-nearest neighbor (KNN) (Rahaman et al., 2019), random forest (RF) (Mohana et al., 2021), and scale-invariant feature transform (SIFT) (Lowe, 2004). However, they have some shortcomings: (a) Classifiers are simple, and the effect of recognition is underperforming. (b) Manual design requires professional knowledge, so it is hard for researchers. Deep learning has rapidly developed in recent years because it can automatically extract features (Lecun et al., 2015) and has achieved excellent performance in vision tasks (Amara et al., 2017). Nevertheless, due to a large number of parameters, designing a good neural network is still a hard task (Suganuma et al., 2017): (a) People made models manually in the early days, which required a lot of professional knowledge (Sun et al., 2020). (b) Neural networks are problem-oriented, whereas manually designed architectures are not. Two main factors affect the performance of neural networks: hyperparameters and training parameters. Training parameters can be learned in the training stage. However, hyperparameters must be set before training. Usually, hyperparameters determine the structure of a network, such as the number and type of layers, kinds of nodes, etc. We hope to find the best hyperparameters for a given dataset in a reasonable time. This process is called hyperparametric optimization. Researchers proved that a gradient descent algorithm is significantly effective for calculating training parameters (Rumelhart et al., 1986). In contrast, there are no explicit strategies for optimizing hyperparameters (Liu et al., 2020).

Most neural networks can be divided into three types: (a) professional knowledge is required for manual design, such as VGG (Ferentinos, 2018) and ResNet (He et al., 2016); (b) semi-automatic design methods like genetic neural networks (Xie and Yuille, 2017), hierarchical evolution (Liu et al., 2017), and others; and (c) fully automatic design, such as when Google introduces the neural architecture search (NAS) concept (Zoph and Le, 2016), which has received considerable attention (Baker et al., 2016; Lu et al., 2018). NAS can search for the best hyperparameters to perform better than manual design. In addition, NAS can reduce trials and errors remarkably. Although NAS is attractive, it still lacks interpretability. Furthermore, model training and verification are costly, and early stopping is meaningful for NAS (Baker et al., 2017; Awad et al., 2021).

This paper studies the hyperparameter optimization of deep learning, and its organization is as follows: (a) Firstly, we examine pertinent technologies. (b) Secondly, the proposed method, as well as the underlying principles, is explained. (c) We then carry out experiments and discuss the results. (d) Finally, conclusions and future research directions are provided.

Related works

Deep learning and convolutional neural networks

Deep learning is a branch of machine learning (Deng and Yu, 2014). It is an automatic feature selection strategy based on neural networks. It can combine low-level features to form abstract high-level features without manual selection. Compared with traditional image recognition and target detection methods, the accuracy and generalization are improved. At present, the main types of neural networks are multilayer perceptrons (MLPs), convolutional neural networks (CNNs), and recurrent neural networks (RNNs), among which CNNs are the most widely used method in classification. Generally, CNNs comprise convolution layers, pooling layers, and fully connected layers. A convolution layer uses correlation information to extract features. A pooling layer (mean pooling or max pooling) compresses the amount of data and parameters, reduces overfitting, and keeps the model invariant to translation, rotation, and scaling. Each neuron (also named a node) in a fully connected layer connects with the previous neurons. Therefore, the multidimensional features are integrated and transformed into several dimensions for classification or detection purposes.

Typical CNN models include AlexNet (Picon et al., 2019), VGGNet, GoogLeNet (Liu et al., 2017), ResNet, MobileNet (Howard et al., 2017), etc. AlexNet is the champion network of the ILSVRC-2012, and it includes five convolution layers and three fully connected layers. According to Bengio et al. (2013), deeper CNNs can extract more representative features. Later, researchers found that blindly increasing the number of layers would slow network convergence (Glorot and Bengio, 2010). Microsoft proposed RESNET with residual blocks and fast connections, which made it possible to build a deeper network (Szegedy et al., 2016). Google proposed MobileNet for mobile and embedded vision applications.

The rapid development of deep learning is inseparable from the extensive use of GPUs. Implementations of CNNs mostly

require GPUs to provide computing support. CNN processes roughly include (a) data preparation and preprocessing; (b) model development, training, and testing; and (c) model deployment. Usually, a dataset can be divided into training, verification, and test sets, with ratios of 7:2:1, 8:1:1, and 6:2:2. Training sets are for learning parameters; verification sets are for optimizing and adjusting hyperparameters; and test sets evaluate performance and generalization. Some public datasets exist, such as PlantVillage (Sm et al., 2020), Kaggle (Ad et al), etc. It is worth noting that many researchers collect their own (Lin et al., 2018; Chen et al., 2020).

Data augmentation

The size and diversity of datasets are essential factors affecting the classification effect of CNNs. Data augmentation can expand the number of images, including moving, flipping, zooming, etc. Deep learning can learn features from images regardless of their positions. Therefore, we can expand datasets through augmentation to avoid overfitting. For example, Perez et al. (Perez and Wang, 2017) developed a new way to use generative adversarial networks (GANs) to make images in different styles.

Neural architecture search

Grid search (GS) is a simple method to find the optimal parameters. However, an exhaustive search may consume time due to the enormous hyperparameter space. Random search (RS) (Andonie and Florea, 2020) explores randomly in the hyperparameter space and improves the performance, but the result may be worse sometimes. For example, the result is not stable. Until now, researchers have proposed many NAS methods: (a) NAS methods based on reinforcement learning (RL); (b) NAS methods based on model optimization; and (c) other improved NAS methods.

(a) NAS methods based on reinforcement learning

Researchers designed a controller to generate strings representing the structures of CNNs, trained each CNN model, and used verification set accuracy as a reward. They optimized the hyperparameters of DCNNs using a novel MARL-based approach (multiagent reinforcement learning) (Iranfar et al., 2022). They then created a multiobjective reward function and applied it to reinforcement learning in order to find the best network with the least latency (Tran and Bae, 2021).

(b) NAS methods based on model optimization

To improve the performance of neural architecture search, researchers proposed ENAs based on evolutionary computing (EC) (Thomas et al) to design CNN architectures. EC is a population-based technology to obtain an optimal global solution. There are some EC-based technologies, such as genetic algorithms (GAs) (Ching-Shih. Deb et al., 2002), particle swarm optimization (PSO) (Kennedy and Eberhart, 1995), and artificial ant colony algorithms (Dorigo et al., 2006). Researchers proposed a two-stage evolutionary search with transfer learning (EvoNAS-TL) (Wen et al., 2021). Also, EPSOCNN, which stands for efficient particle swarm optimization, is suggested as a way to improve CNN architectures (Wang et al., 2020).

(c) Other improved NAS methods

To limit the search space, Yu et al. (2021) and Sun et al. (2019) proposed block-based methods. However, the results are insufficient and unstable due to the lack of theoretical support. Hu et al. (2021) proposed a new performance estimation metric named random-weight evaluation (RWE) to quantify the quality of CNNs. Lu et al. (2018) proposed NSGANet, an evolutionary algorithm that combines prior knowledge from handcrafted architectures with an exploration comprising crossover and mutation. Some software packages provide search functions, such as pyGPGO and Optunity (Bergstra et al., 2011), Hyperopt–Sklearn (Bergstra et al., 2015), etc.

Many CNN models are challenging to apply on mobile/edge devices due to limited resources such as memory capacity and power consumption. Researchers have carefully designed some lightweight networks. Donegan et al. (2021) used a differentiable NAS method to find efficient CNNs for the Intel Movidius Vision Processing Unit (VPU), achieving state-of-the-art accuracy on ImageNet. An FPGA-based CNN accelerator (field programmable gate array) was proposed (Fan et al., 2020) with an accurate performance model of hardware design. Intelligent edge-cloud scenarios are expected to meet diverse requirements.

All the above studies do not effectively record evolutionary information, so they cannot guide the whole search process based on experience. In contrast, Bayesian optimization (Frazier, 2018) assumes the search space as a Gaussian distribution, learns experience in search processes, and calculates better parameters iteratively (Wistuba; Gupta et al., 2017; Ji et al., 2019; Basha et al., 2021). However, these methods still cost enormous computing resources. As a result, this paper suggests a heuristic target-dependent method that only needs small samples and is entirely automatic.

Proposed method

From the above, we know that neural architecture search is time-consuming and requires many resources. Therefore, we intend to optimize the hyperparameter exploration process. In short, the main contributions of this paper are as follows:

a. Method: We proposed a target-dependent search method that only needs small samples. Besides accuracy, we use precision and recall to promote generalization. Also, our method can find searches that are not working well and stop them early to save time and resources.
b. Dataset: We collected eight kinds of rapeseed images and created the dataset RSDS.
c. Comparison: Horizontally, we compare each TD-NAS based on VGGNets. Vertically, we explore the TD-NAS based on VGGNet-D (VGGNet-16). Furthermore, we test our method on two additional open datasets, Pl@ntNet and ICL-Leaf.

Infrastructure and hyperparameters

As for the primary network architecture, we chose VGGNet, a typical convolutional network with six deepening structures labeled A, A-LRN, B, C, D, and E. Here we select A, B, D, and E for experiments, and D is the most famous model named VGGNet-16 (Figure 1).

We keep the number and position of convolution and pooling layers fixed, while the layer number, the dropout rate, and the neuron number of fully connected layers can be changed. Table 1 shows the hyperparameter space. It is worth noting that there is at least one fully connected layer, and the number of neurons in the last fully connected layer is eight to produce the output—eight classes of rapeseed.

Search principle and Bayesian optimization

The aim is to find the hyperparameters of a model with the best performance on verification sets. Let T be the objective function for getting maximum accuracy (ACC).

$$hp^{*} = \underset{hp \in D}{\operatorname{argmax}} T(hp) \tag{1}$$

In formula (1), D is a hyperparameter space. We can create a model for each hp in D, train the model, and evaluate its performance on verification sets. This paper separates the RSDS dataset into three parts: the training set, the validation set, and the test set, with a ratio of 7:2:1. We use formula (2) as the judgment criteria (jc) for evaluating model qualities:

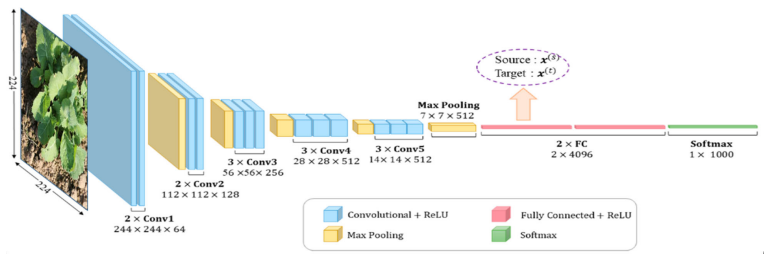


FIGURE 1
VGGNet-16.

TABLE 1 Hyperparameter space.

Level	Layer	Name	Values	Type
Architectural	*FC*	*LN*	{1, 2, 3}	Integer
Internal	*FC*	*DR*	{0.0, 0.1, 0.2, 0.3, 0.4, 0.5, 0.6, 0.7, 0.8, 0.9}	Float
Internal	*FC*	*NN*	{512, 1,024, 2,048, 4,096}	Integer
External	*OU*	ω_2 = 1- ω_1	{1.0, 0.9, 0.8, 0.7, 0.6, 0.5}	Float
External	*OU*	β	{0.5, 1.0, 2.0}	Float

FC, fully connected; LN, layer number; DR, dropout rate; NN, neuron number; OU, output. ω, β: parameters in formula (2) (discussed in the next section.)

$$jc = \omega_1 Acc + \omega_2 F_\beta \quad (2)$$

Here, F_β (balanced *F*-score) is the harmonic average of precision and recall. Usually, the smaller the F-score, the better the generalization. Therefore, we refine formula (1) to formula (3):

$$hp^* = \underset{hp \in D}{\operatorname{argmax}} T(hp, \quad jc) \quad (3)$$

The following Algorithm 1 gives the naive hyperparameter search process:

Input: Space_D, Data_T, Data_VOutput: hp*(1) Get the training set (Data_T) and the verification set (Data_V).(2) Select a judgment criterion: *ACC*.(3) Set hyperparameter search space (Space_D) and initialize one hyperparameter (hp):(4) Generate a CNN model with the hp.(5) Train the model, and verify it.(6) Select the next hyperparameter (hp), repeat (4), or Quit.(7) Output the optimal hyperparameter (hp*).

ALGORITHM 1
HPS: Hyperparameters Search.

Algorithm 1 does not record evolutionary information, so it cannot guide search processes effectively. Even if adopting a random search, uncertainty still exists.

We propose a heuristic target-dependent search method. The so-called heuristic means our approach can evaluate a better location and start a new search. Here, we choose Bayesian optimization, which assumes the superparameter space as a Gaussian distribution and obtains better candidates each time. Our method introduces a stop criterion to reduce the search scale without lowering generalization. The so-called target-dependent means that the explored architecture is not universal and only valid for specific crops. We can quickly rerun the proposed method to search out new architectures when facing new species.

Bayesian optimization has two components: (a) Bayesian statistics for constructing an objective function (typically a Gaussian process); and (b) acquisition function for calculating the following sampling points. After initializing several points, Bayesian optimization can calculate *a posteriori* and iterate reasoning until meeting an exit condition. Algorithm 2 shows our TD-NAS method based on Bayesian optimization, where GP is a Gaussian process, Acq_F is an acquisition function, and Dyn_QF is a dynamic quit function. In Algorithm 2, we focus on steps (6), (4), and (3). Step (4) costs massive resources for VGGNet training and verifying, so step (6) should select hyperparameters elegantly to reduce the number of models. Furthermore, step (4) should stop the training and verifying processes when there are unpromising detections. Step (3) checks the dynamic quit conditions and decides whether to quit or not.

Input: VGGNet, Space_D, Data_T, Data_V, GP, Acq_F, Dyn_QFOutput: x*(1) Init S = {(x_i, y_i)}, y_i = f(x_i), $x_i \in$ *Space_D*, let f~GP(μ, K).(2) Select a judgment criterion: *jc*.(3) While not Dyn_QF() do:(4) Train and verify VGGNet(x, Data_T, Data_V) with unpromising detections.(5) Calculate p(y | x, S).(6) Acq_F(x, p(y | x, S)), get x_{new}.(7) y_{new} = f(x_{new}).(8) S = S U (x_{new}, y_{new}).(9) Output the optimal hyperparameter (x*).

ALGORITHM 2
TD-NAS: Target-dependent neural architecture search.
HPS: Hyperparameters Search.

Step (6): The acquisition function strikes a balance between exploration and exploitation.

In Bayesian optimization, the acquisition function (Acq_F) is critical for generating points according to prior knowledge. Exploitation means evaluating at expected points because global optima are likely to reside there. Exploration means considering uncertain points is helpful because objects tend to be far from where we have measured them. Usually, there are three typical acquisition functions: expected improvement (EI), entropy search (ES), and knowledge gradient (KG). The expected value of EI is easy to figure out, which makes it a popular acquisition function.

Let $f_n^* = \max_{m \le n} f(x_m)$ be the max previous value. We have one other position, x, to be evaluated, and then we get f (x). Now, the best-observed point is either *f(x)* or f_n^*. The improvement is then $f(x) - f_n^*$; if this quantity is positive, else 0, mark as $[\mathrm{f}(\mathrm{x}) - f_n^*]^+$ for convenience. Unfortunately, we should train and validate the entire network to get *f(x)*. Instead, we can take the expected value of this improvement and define formula (4):

$$x_{n+1} = \operatorname{argmax} \mathrm{EI}_n(x) \quad (4)$$

Here, $\mathrm{EI}_n(x) = E_n[[\mathrm{f}(\mathrm{x}) - f_n^*]^+]$, and E_n indicates that the expectation is taken under the posterior distribution, as shown in formula (5): *f(x)* given $x_{1:n}$, $y_{1:n}$ is normally distributed with mean μ_n (x) and variance σ_n^2(x).

$$f(x) \,|\, f(x_{1:n}) \sim \mathrm{Normal}\left(\mu_n(x), \sigma_n^2(x)\right) \quad (5)$$

Unlike the *f(x)* in step 4 of Algorithm 2, EI_n (*x*) is low cost to observe and allows for easy evaluation of first- and second-order derivatives, as shown in formula (6):

$$\mathrm{EI}_n(x) = \Delta_n(\mathrm{x})\Phi\left(\frac{\Delta_n(\mathrm{x})}{\sigma_n(\mathrm{x})}\right) + \sigma_n(\mathrm{x})\varphi\left(\frac{\Delta_n(\mathrm{x})}{\sigma_n(\mathrm{x})}\right) \quad (6)$$

The definitions of $\Phi()$,$\varphi()$ can be found in (Jones et al., 1998). Here, $\Delta_n(\mathrm{x}) = \mu_n(\mathrm{x}) - f_n^*$ is the expected difference between the proposed point x and the previous best. Note that $\mathrm{EI}_n(x)$ balances between high expected quality ($\Delta_n(x)$) versus high uncertainty ($\sigma_n(x)$).

Step (4): Training and validating the VGGNet without unpromising detections.

We calculate verification errors when training and verifying. If the error exceeds the average prior value, it is unpromising to go further. Figure 2 shows five hyperparameters in search, including two unpromising detections. Stopping these unpromising detections early can save resources and time.

Step (3): Making dynamic quit decisions.

Figure 3 gives dynamic quit conditions and their generation approach. To control the search process, a dynamic quit function (Dyn_QF) uses these conditions, including whether the queue of the hyperparameter space is empty or the maximum number of iterations has been reached.

Experiment and discussion

Dataset and experimental condition

We took photos using a Canon EoS6D camera, which has 20.2 million effective pixels and a maximum resolution of 5,472 × 3,648. We then resized each image to 224 × 224 pixels to improve the processing speed. To run programs, we used an HP-

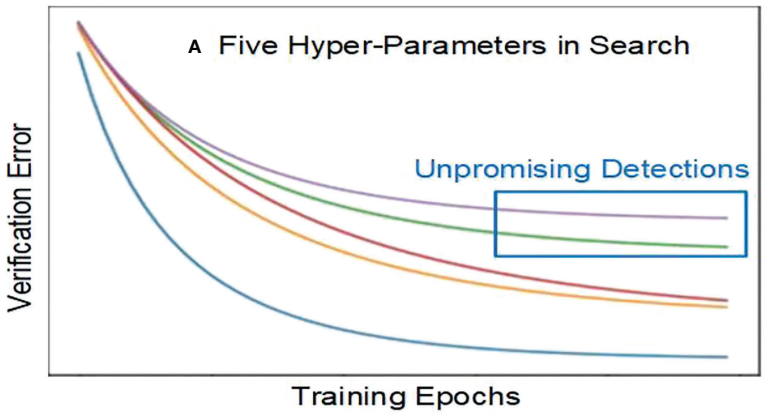


FIGURE 2
Unpromising detections.

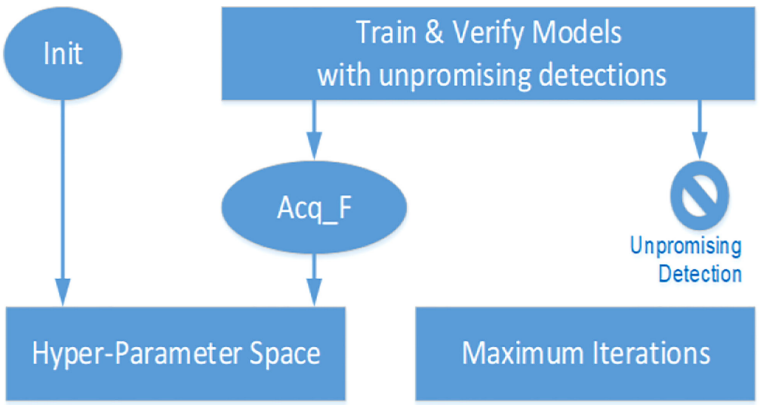


FIGURE 3
Dynamic quit conditions. Generation approach.

OMEN laptop with an i9-9880H CPU, 32 GB of memory, an NVIDIA RTX2080 graphics card (8 GB), and Python installed.

We collected rapeseed images at an experimental station of Anhui Agricultural University, located at 117.2° east longitude and 31.5° north latitude, in Sanhe town, Hefei, China. We obtained eight kinds of rapeseed images, at least 1,000 of each class ($C_i|i$ = 1,…, 8), as shown in Figure 4, named RSDS. We divided the RSDS into three parts: training set (Tr), verification set (V), and test set (Te), with a ratio of 7:2:1. We randomly obtained *RSDS-0.1K* with (Tr, V, Te = 700, 200, 100) images per class. Using data augmentation, we obtained more sets as follows: *RSDS-0.2K* (Tr, V, Te = 1,400, 400, 200), *RSDS-0.4K* (Tr, V, Te = 2,800, 800, 400), and *RSDS-1.0K* (Tr, V, Te = 7,000, 2,000, 1,000).

Result discussion

We use VGGNet as the base framework. When training, the initial learning rate is 0.01, and the epoch size is 50. For comparing two algorithms, "better" means (a) fewer attempts for the same score and (b) a higher score after the same number of tries.

For horizontal comparisons, (a) we set $\omega_1 = 1$, which means only accuracy is the evaluation indicator. In Table 2, the verification accuracy of TD-NAS based on VGGNet-16 reaches 81.38%. However, the result obtained on VGGNet-E (VGGNet-19) is worse than the original, indicating that Bayesian optimization also has limitations in dealing with deep networks. The number of neurons in the last fully connected layer fixes eight to output the probability values of rapeseed classes through a Softmax function. (b) Do not fix ω_1. Instead, use *jc* as the indicator in formula (2) (Table 3).

Keep the fully connected parameters from Table 2 unchanged and take into account ω_1 and β. After searching, we still get the highest verification accuracy (81.06%, Table 3) based on VGGNet-D, which is slightly lower than the accuracy of VGGNet-16 (81.38%, Table 2). However, the verification accuracy of TD-NAS based on VGGNet-E has increased from 78.44% (Table 2) to 79.94% (Table 3).

Now, for vertical comparisons, (a) we use four datasets (Table 4) to search for TD-NAS on VGGNet-D (VGGNet-16).

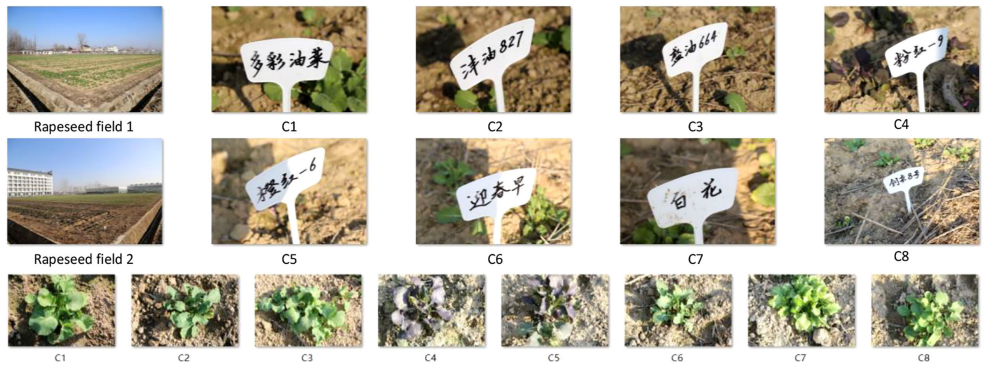


FIGURE 4
Rapeseed dataset (RSDS).

TABLE 2 TD-NAS and original VGGNet (A, B, D, and E) (*RSDS-0.1K*, $\omega_1 = 1$).

	VGGNet-A		VGGNet-B		VGGNet-D		VGGNet-E	
	Original	**TD-NAS**	**Original**	**TD-NAS**	**Original**	**TD-NAS**	**Original**	**TD-NAS**
V-ACC	74.06%	78.81%	77.38%	77.94%	79.19%	**81.38%**	80.75%	78.44%
FC								
LN	3	3	3	3	3	3	3	3
DR	{0.7, 0.7, 0.0}	{0.6, 0.4, 0.0}	{0.7, 0.7, 0.0}	{0.7, 0.5, 0.0}	{0.7, 0.7, 0.0}	{0.6, 0.5, 0.0}	{0.7, 0.7, 0.0}	{0.5, 0.3, 0.0}
NN	{4,096, 4,096, 8}	{2,048, 2,048, 8}	{4,096, 4,096, 8}	{2,048, 2,048, 8}	{4,096, 4,096, 8}	{4,096, 2,048, 8}	{4,096, 4,096, 8}	{2,048, 2,048, 8}
OU								
ω_1	1	1	1	1	1	1	1	1
β	–	–	–	–	–	–	–	–

V-ACC, verification accuracy; FC, fully connected; LN, layer number; DR, dropout rate; NN, neuron number; OU, output. Bold means To highlight the biggest verification accuracy (V-ACC).
Bold values mean to highlight the biggest verification accuracy (V-ACC).

(b) We choose the model made by *RSDS-0.1K*, figure out how accurate it is on all of the test sets, and compare it to other models (Table 5). In Table 4, it can be seen that the bigger the dataset size, the higher the verification accuracy (91.23%, Table 4). It is worth noting that from *RSDS-0.4K* to *RSDS-1.0K*, the promotion of verification accuracy is only 1.95%, but the amount of training data has increased by 1.5 times. Meanwhile, the time consumptions of *RSDS-0.1K*, *RSDS-0.2K*, *RSDS-0.4K*, and *RSDS-1.0K* are approximately 00:42:32, 01:11:36, 04:20:38, and 07:26:00, respectively. The costs of training and verification vary greatly, but the benefits are limited.

The TD-NAS mentioned above on VGGNet-D (VGGNet-16) searches each dataset to generate models. Table 5 shows the accuracy of these models on test sets; the diagonal part of the table shows the accuracy of each model on its own test set, whereas the first column is the accuracy of the model generated by *RSDS-0.1K* on all test sets. It is worth noting that the test accuracy of the model trained on small samples is not much different from that trained on large ones.

Figure 5 shows four confusion matrixes generated on each test set separately, and the model is TD-NAS on VGGNet-D (VGGNet-16). Overall, our model performed well across different test sets. It is worth noting that the two pairs [C4, C5] and [C7, C8] are far more likely than others to misjudge each other. When a person looks at the rapeseed images of these two pairs, they look very similar in color and shape.

Pl@ntNet-300K and ICL-Leaf datasets

A novel image dataset with high intrinsic ambiguity was presented (Camille et al., 2021), built explicitly for evaluating and comparing set-valued classifiers. It consists of 306,146 images covering 1,081 species, with two particular features: (a) The dataset has a strong class imbalance, which means that a few species account for most images. (b) Many species are visually similar (Figure 6), making identification difficult even for eye experts.

Table 6 shows that test accuracy depends strongly on the number of images per class. We selected eight species, including *Cirsium arvense*, and searched for architecture based on VGG-16. Let [s1, s2] be the interval of the image number per class (for example (Rahaman et al., 2019; Sun et al., 2019),), and we searched twice with s1 and s2 per class (for >2,000, set [2,500,

TABLE 3 TD-NAS (*RSDS-0.1K*, ω_1 not fixed).

	TD-NAS			
	VGGNet-A	**VGGNet-B**	**VGGNet-D**	**VGGNet-E**
V-ACC	77.56%	75.13%	**81.06%**	79.94%
FC				
LN	3	3	3	3
DR	{0.6, 0.4, 0.0}	{0.7, 0.5, 0.0}	{0.6, 0.5, 0.0}	{0.5, 0.3, 0.0}
NN	{2,048, 2,048, 8}	{2,048, 2,048, 8}	{4,096, 2,048, 8}	{2,048, 2,048, 8}
OU				
ω_1	0.9	0.9	0.8	0.9
β	2	0.5	1	0.5

V-ACC, verification accuracy; FC, fully connected; LN, layer number; DR, dropout rate; NN, neuron number; OU, output. Bold means To highlight the biggest verification accuracy (V-ACC). Bold values mean to highlight the biggest verification accuracy (V-ACC).

TABLE 4 TD-NAS on VGGNet-D (VGGNet-16) (ω_1 *not fixed*).

	TD-NAS on VGGNet-D (VGGNet-16)			
	RSDS-0.1K	***RSDS-0.2K***	***RSDS-0.4K***	***RSDS-1.0K***
V-ACC	81.06%	86.53%	89.28%	**91.23%**
FC				
LN	3	3	3	3
DR	{0.6, 0.5, 0.0}	{0.7, 0.3, 0.0}	{0.4, 0.6, 0.0}	{0.5, 0.4, 0.0}
NN	{4,096, 2,048, 8}	{4,096, 2,048, 8}	{2,048, 2,048, 8}	{2,048, 1,024, 8}
OU				
ω_1	0.8	0.9	0.9	0.8
β	1	0.5	1	2

V-ACC, verification accuracy; FC, fully connected; LN, layer number; DR, dropout rate; NN, neuron number; OU, output. Bold means To highlight the biggest verification accuracy (V-ACC). Bold values mean to highlight the biggest verification accuracy (V-ACC).

TABLE 5 Test accuracy comparison.

T-ACC	TD-NAS on VGGNet-D (VGGNet-16)			
	RSDS-0.1K	RSDS-0.2K	RSDS-0.4K	RSDS-1.0K
RSDS-0.1K	81.13%	–	–	–
RSDS-0.2K	79.94%	81.44%	–	–
RSDS-0.4K	80.22%	–	83.13%	–
RSDS-1.0K	81.68%	–	–	82.73%

T-ACC, test accuracy.

3,000]), then calculated the mean accuracy. From Table 6, we found that our method got higher accuracy than Ref (Camille et al., 2021). claimed, except for the situation ">2,000". The results in Ref (Camille et al., 2021). were made with ResNet50, which has 49 convolutional layers and one fully connected layer. This architecture is deeper than ours and better at handling large samples.

Another public leaf dataset called the ICL (Hu et al., 2012) was built by the Intelligent Computing Laboratory (ICL) at the Institute of Intelligent Machines, Chinese Academy of Sciences. It contains 16,851 samples from 220 species, with 26 to 1,078 samples per species (Figure 7).

We selected eight species, including *Amorpha fruticosa*, and searched for architecture based on VGG-16 two times. We got an average test accuracy of 92.31%, a little more than 92.08% in Ref (Xiao et al., 2010)., which used a traditional machine-learning method named HOG-MMC (orientation histogram based on dimension reduction of maximum edge criterion).

From the above, we proved that our method could quickly discover new architectures when faced with Pl@ntNet-300K and ICL-Leaf datasets. The contribution of our method includes three features: (a) small sample sizes, (b) stable generalization, and (c) free of unpromising detections. In experiments of Pl@ntNet-300K and ICL-Leaf, as for feature a, we selected [500, 2,000] and [1,000, 1,200]

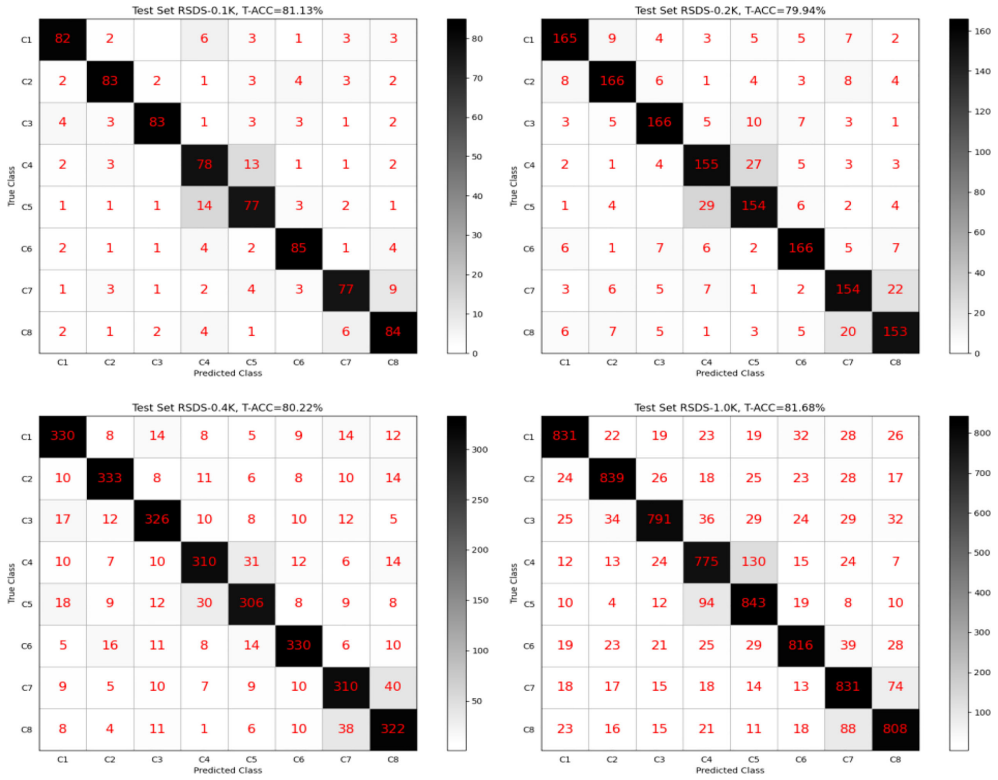


FIGURE 5
Confusion matrixes generated by TD-NAS on VGGNet-D (VGGNet-16).

FIGURE 6
Examples of visually similar images belonging to two different classes.

TABLE 6 Test accuracy comparison.

Number of images per class	Mean accuracy claimed in (Camille et al., 2021)	Ours, TD-NAS on VGGNet-D (VGGNet-16)
10–50	35%	41.25%
50–500	59%	63.38%
500–2,000	79%	80.97%
>2,000	93%	86.75%

images per class and got average test accuracies represented above. As for feature b, we increased the number of images per class to [2,500, 3,000] and [1,500, 2,000], and test accuracy had risen a little (< 6%), but time cost ascended (> 26%). As for feature c, we set a switch to control NAS with or without unpromising detections, and statistics showed that more than 19% of detections are unpromising. To sum up, due to features a and c, we can quickly find out new architectures when faced with new species but still get feature b's stable generalization.

Conclusion

This paper aims to design a novel target-dependent neural architecture search method based on VGGNet. This goal was successfully achieved on a self-built dataset with eight kinds of rapeseed images. We select accuracy, precision, and recall as the evaluation indicators. We adopt Bayesian optimization to obtain better candidate parameters and introduce a stop criterion for optimizing the dynamic search process. Results show that the

FIGURE 7
Typical samples for each species.

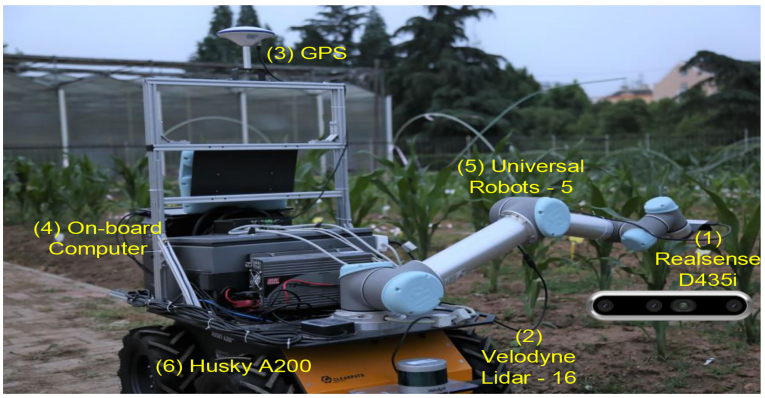


FIGURE 8
Our robot platform (capture pictures by (1) and run models on (4)).

test accuracy of the model trained on small samples is not much different from those trained on large ones. The generalization of the model generated by our method is not sensitive to dataset size, making it meaningful to search out models from small samples when facing new objects. For example, we tested our method on two other open datasets, Pl@ntNet and ICL-Leaf.

Due to the enormous model structure and parameter space, this paper only dealt with fully connected layers and the output layer. We kept the original network structure unchanged, such as the kernel and pooling size in convolutional layers. However, the full network structure search, including convolutional layers, is universal without margin. Some literature has pointed out that Bayesian optimization is only suitable for medium-sized problems (Shahriari et al., 2015). When faced with new objects, how to quickly search a minimized network structure is still an attractive topic. Figure 8 shows our robot platform, with our models running on (4), an onboard computer, and real-time images captured by (1), a Realsense D435i camera. We continue to explore in-depth research questions, such as changing kernels and pooling sizes in convolutional layers (Franchini et al., 2022). We are thinking about using the proposed method on more CNN frameworks in the future.

Data availability statement

The self-build dataset RSDS will be made available by the authors. Another two open datasets (Pl@ntNet and ICL-Leaf) can be found on the official sites or the Internet, such as: https://zenodo.org/record/5645731.

Author contributions

LYF and YR planned and designed the research. JXL, JT, and QLH performed the experiments. LYF wrote the manuscript. SWL supervised this work and reviewed the manuscript. All authors contributed to the article and approved the submitted version.

Funding

948 Project of Introduction of International Advanced Agricultural Science and Technology by Ministry of Agriculture, Ministry of Agriculture, China, No.2016-X34.

Conflict of interest

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

Publisher's note

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

References

Ad, A., De, B., and Aeh, B. (2020). An optimized model based on convolutional neural networks and orthogonal learning particle swarm optimization algorithm for plant diseases diagnosis. *Swarm Evol. Computation* 52, 100616. doi: 10.1016/j.swevo.2019.100616

Amara, J., Bouaziz, B., and Algergawy, A. (2017). A Deep Learning-based Approach for Banana Leaf Diseases Classification. Datenbanksysteme für Business, Technologie und Web (BTW 2017), 17 Fachtagung des GI-Fachbereichs,Datenbanken und Informationssysteme" (DBIS), 6-10 März 2017, Stuttgart, Germany, Workshopband.

Andonie, R., and Florea, A. C. (2020). Weighted random search for CNN hyperparameter optimization. *Int. J. Computers Commun. Control (IJCCC)* 15 (2), 3868. doi: 10.15837/ijccc.2020.2.3868

Awad, N., Mallik, N., and Hutter, F. (2021). DEHB: Evolutionary hyberband for scalable, robust and efficient hyperparameter optimization. In *Proceedings of the Thirtieth International Joint Conference on Artificial Intelligence*, Main Track, 2147–2153. doi: 10.24963/ijcai.2021/296

Baker, B., Gupta, O., Naik, N., and Raskar, R. (2016). Designing neural network architectures using reinforcement learning. In *5th International Conference on Learning Representations, 2017*. doi: 10.48550/arXiv.1611.02167

Baker, B., Gupta, O., Raskar, R., Naik, N., et al. (2017). Accelerating neural architecture search using performance prediction. In *International Conference on Learning Representations (ICLR) Workshops 2018*. doi: 10.48550/arXiv.1705.10823

Barré, P., Durand, H., Chenu, C., Meunier, P., Montagne, D., Castel, G., et al. (2017). Geological control of soil organic carbon and nitrogen stocks at the landscape scale. *Geoderma* 285, 50–56. doi: 10.1016/j.geoderma.2016.09.029

Basha, S., Vinakota, S. K., Dubey, S. R., Pulabaigari, V., and Mukherjee, S. (2021). AutoFCL: automatically tuning fully connected layers for handling small dataset. *Neural Computing App.* 33, 8055–8065. doi: 10.1007/s00521-020-05549-4

Bengio, Y., Courville, A., and Vincent, P. (2013). Representation learning: A review and new perspectives. *IEEE Trans. Pattern Anal. Mach. Intelligence* 35 (8), 1798–1828. doi: 10.1109/TPAMI.2013.50

Bergstra, J., Bardenet, R., Bengio, Y., and Kégl, B. (2011). "Algorithms for hyper-parameter optimization," in *Proceedings of the 24th International Conference on Neural Information Processing Systems; Granada, Spain: Curran Associates Inc.* (Red Hook, NY, USA: Curran Associates Inc.) 2546–2554. doi: 10.5555/2986459.2986743

Bergstra, J., Komer, B., Eliasmith, C., Yamins, D., and Cox, D. D. (2015). Hyperopt: a Python library for model selection and hyperparameter optimization. *Comput. Sci. Discov* 8 (1). doi: 10.1088/1749-4699/8/1/014008

Camille, G., Alexis, J., Pierre, B., Antoine, A., Jean-Christophe, L., Mathias, C., et al. (2021). "Pl@ntNet-300K: a plant image dataset with high label ambiguity and a long-tailed distribution," in *Proc. of Thirty-fifth Conference on Neural Information Processing Systems, Datasets and Benchmarks Track*. doi: 10.5281/zenodo.5645731

Chen, J., Chen, J., Zhang, D., Sun, Y., and Nanehkaran, Y. A. (2020). Using deep transfer learning for image-based plant disease identification. *Comput. Electron. Agricult.* 173, 105393. doi: 10.1016/j.compag.2020.105393

Ching-Shih. Deb, K., Pratap, A., Agarwal, S., and Meyarivan, T. (2002). A fast and elitist multiobjective genetic algorithm: NSGA-II. *IEEE Trans. Evol. Comput.* 6 (2), 182–197. doi: 10.1109/4235.996017

Deng, L., and Yu, D. (2014). Deep learning: Methods and applications. *Foundations Trends Signal Processing* 7 (3), 197–387. doi: 10.1561/2000000039

Donegan, C., Yous, H., Sinha, S., and Byrne, J. (2021). "VPU specific CNNs through neural architecture search," in *2020 25th International Conference on Pattern Recognition (ICPR)*, 2021 pp. 9772–9779. doi: 10.1109/ICPR48806.2021.9412794

Dorigo, M., Birattari, M., and Stützle, T. (2006). Ant colony optimization. *IEEE Comput. Intell. Magazine* 1 (4), 28–39. doi: 10.1109/MCI.2006.329691

Fan, H., Ferianc, M., Liu, S., Que, Z., Niu, X., and Luk, W. (2020). "Optimizing FPGA-based CNN accelerator using differentiable neural architecture search," in *2020 IEEE 38th International Conference on Computer Design (ICCD)*, Hartford, CT. 465–468. doi: 10.1109/ICCD50377.2020.00085

Ferentinos, K. P. (2018). Deep learning models for plant disease detection and diagnosis. *Comput. Electron. Agricult.* 145, 311–318. doi: 10.1016/j.compag.2018.01.009

Franchini, G., Ruggiero, V., Porta, F., and Zanni, L. (2022). Neural architecture search *via* standard machine learning methodologies. *Mathematics Eng.* 5 (1), 1–21. doi: 10.3934/mine.2023012

Frazier, P. I. (2018). A tutorial on Bayesian optimization. *INFORMS TutORials in Operations Research*, 255–278. doi: 10.1287/educ.2018.0188

Fu, R., Li, B., Gao, Y., and Ping, W. (2016). "Content-based image retrieval based on CNN and SVM," in *2016 2nd IEEE International Conference on Computer and Communications (ICCC)*. 2016, pp. 638–642. doi: 10.1109/CompComm.2016.7924779

Glorot, X., and Bengio, Y. (2010). Understanding the difficulty of training deep feedforward neural networks. *J. Mach. Learn. Res.* 9, 249–256.

Gupta, A., Ong, Y.-S., and Feng, L. (2017). "Insights on transfer optimization: Because experience is the best teacher," in *IEEE Transactions on Emerging Topics in Computational Intelligence*, Vol. 2, 51–64. doi: 10.1109/TETCI.2017.2769104

He, K., Zhang, X., Ren, S., and Sun, J. (2016). Deep residual learning for image recognition. *IEEE* 2016, 770–778. doi: 10.1109/CVPR.2016.90

Howard, A. G., Zhu, M., Chen, B., Kalenichenko, D., Wang, W., Weyand, T., et al. (2017). MobileNets: Efficient convolutional neural networks for mobile vision applications. *arXiv preprint*. doi: 10.48550/arXiv.1704.04861

Hu, S., Cheng, R., He, C., Lu, Z., Wang, J., Zhang, M., et al. (2021). Accelerating multi-objective neural architecture search by random-weight evaluation. *Complex Intelligent Syst* 2021. doi: 10.1007/s40747-021-00594-5

Hu, R., Jia, W., Ling, H., and Huang, D. (2012). "Multiscale distance matrix for fast plant leaf recognition," in *IEEE Trans. Image Processing*, Vol. 21. 4667–4672.

Iranfar, A., Zapater, M., and Atienza, D. (2022). Multiagent reinforcement learning for hyperparameter optimization of convolutional neural networks. *IEEE Trans. Computer-Aided Design Integrated Circuits Syst.* 41, 1034–1047. doi: 10.1109/tcad.2021.3077193

Ji, D., Jiang, Y., Qian, P., and Wang, S. (2019). "A novel doubly reweighting multisource transfer learning framework," in *IEEE Transactions on Emerging Topics in Computational Intelligence*. 3, 380–391. doi: 10.1109/TETCI.2018.286832

Jones, D. R., Schonlau, M., and Welch, W. J. (1998). Efficient global optimization of expensive black-box functions. *J. Global Optimization* 13, 455–492. doi: 10.1023/A:1008306431147

Kennedy, J., and Eberhart, R. (1995). "Particle swarm optimization," in *Icnn95-international Conference on Neural Networks*. vol.4, pp. 1942–1948. doi: 10.1109/ICNN.1995.488968

Lecun, Y., Bengio, Y., and Hinton, G. (2015). Deep learning. *Nature* 521 (7553), 436. doi: 10.1038/nature14539

Lee, S., and Chang, J. H. (2017). Deep learning ensemble with asymptotic techniques for oscillometric blood pressure estimation. *Comput. Methods Programs Biomed.* 151, 1–13. doi: 10.1016/j.cmpb.2017.08.005

Li, Z., Guo, R., Li, M., Chen, Y., and Li, G. (2020). A review of computer vision technologies for plant phenotyping. *Comput. Electron. Agricult.* 176, 105672. doi: 10.1016/j.compag.2020.105672

Li, Y., Nie, J., and Chao, X. (2020). Do we really need deep CNN for plant diseases identification? *Comput. Electron. Agric.* 178 (3), 105803. doi: 10.1016/j.compag.2020.105803

Lin, Z., Mu, S., Shi, A., Pang, C., and Sun, X. (2018). A novel method of maize leaf disease image identification based on a multichannel convolutional neural network. *Transactions of the ASABE* 61 (5), 1461–1474. doi: 10.13031/trans.12440

Liu, H., Simonyan, K., Vinyals, O., Fernando, C., and Kavukcuoglu, K. (2017). Hierarchical representations for efficient architecture search. *ICLR 2018*. doi: 10.48550/arXiv.1711.00436

Liu, Y., Sun, Y., Xue, B., Zhang, M., and Yen, G. (2020). A survey on evolutionary neural architecture search. *IEEE Trans Neural Netw Learn Syst.*, 1–21. doi: 10.1109/TNNLS.2021.3100554

Liu, B., Zhang, Y., He, D. J., and Li, Y. (2017). Identification of apple leaf diseases based on deep convolutional neural networks. *Symmetry* 10 (1), 11. doi: 10.3390/sym10010011

Lowe, D. G. (2004). Distinctive image features from scale-invariant keypoints. *Int. J. Comput. Vision* 60 (2), 91–110. doi: 10.1023/B:VISI.0000029664.99615.94

Lu, Z., Whalen, I., Boddeti, V., Dhebar, Y., Deb, K., Goodman, E., et al. (2018). NSGA-net: Neural architecture search using multi-objective genetic algorithm. *GECCO 2019*. doi: 10.48550/arXiv.1810.03522

Mohana, R. M., Reddy, C., Anisha, P. R., and Murthy, B. (2021). Random forest algorithms for the classification of tree-based ensemble. *Mater. Today: Proc.* 12, In Press. doi: 10.1016/j.matpr.2021.01.788

Pawara, P., Okafor, E., Surinta, O., Schomaker, L., and Wiering, M. (2017). "Comparing local descriptors and bags of visual words to deep convolutional neural networks for plant recognition," in *Proceedings of the 6th International Conference on Pattern Recognition Applications and Methods-ICPRAM*, 479–486. doi: 10.5220/0006196204790486

Perez, L., and Wang, J. (2017). The effectiveness of data augmentation in image classification using deep learning.*arXiv preprint*. doi: 10.48550/arXiv.1712.04621

Picon, A., Alvarez-Gila, A., Seitz, M., Ortiz-Barredo, A., Echazarra, J., Johannes, A., et al. (2019). Deep convolutional neural networks for mobile capture device-based crop disease classification in the wild. *Comput. Electron. Agricult.* 161, 280–290. doi: 10.1016/j.compag.2018.04.002

Purohit, S., Viroja, R., Gandhi, S., and Chaudhary, N. (2016). Automatic plant species recognition technique using machine learning approaches. *Int. Conf. Computing Netw. Commun* 2015, 710–719. doi: 10.1109/CoCoNet.2015.7411268

Rahaman, M. A., Hossain, E., and Hossain, M. F. (2019). "A color and texture based approach for the detection and classification of plant leaf disease using KNN classifier," in *2019 International Conference on Electrical, Computer and Communication Engineering (ECCE)*. pp. 1–6. doi: 10.1109/ECACE.2019.8679247

Rumelhart, D., Hinton, G. E., and Williams, R. J. (1986). Learning representations by back propagating errors. *Nature* 323 (6088), 533–536. doi: 10.1038/323533a0

Rumpf, T., Mahlein, A. K., Steiner, U., Oerke, E. C., Dehne, H. W., and Plümer, L. (2010). Early detection and classification of plant diseases with support vector machines based on hyperspectral reflectance. *Comput. Electron. Agricult.* 74 (1), 91–99. doi: 10.1016/j.compag.2010.06.009

Shahriari, B., Swersky, K., Wang, Z., Adams, R. P., and Freitas, N. D. (2015). Taking the human out of the loop: A review of Bayesian optimization. *Proc. IEEE* 104 (1), 148–175. doi: 10.1109/JPROC.2015.2494218

Sm, A., Rs, A., and Dr, A. (2020). Deep convolutional neural network based detection system for real-time corn plant disease recognition. *Proc. Comput. Sci.* 167, 2003–2010. doi: 10.1016/j.procs.2020.03.236

Suganuma, M., Shirakawa, S., and Nagao, T. (2017). A genetic programming approach to designing convolutional neural network architectures. *GECCO '17: Proceedings of the Genetic and Evolutionary Computation Conference*, 497–504. doi: 10.1145/3071178.3071229

Sun, Y., Xue, B., Zhang, M., and Yen, G. G. (2019). Completely automated CNN architecture design based on blocks. *IEEE Trans. Neural Networks Learn. Systems* 99, 1–13. doi: 10.1109/TNNLS.2019.2919608

Sun, Y., Xue, B., Zhang, M., and Yen, G. G. (2020). Evolving deep convolutional neural networks for image classification. *IEEE Trans. Evol. Computation* 24 (2), 394–407. doi: 10.1109/TEVC.2019.2916183

Szegedy, C., Ioffe, S., Vanhoucke, V., and Alemi, A. (2016). Inception-v4, inception-ResNet and the impact of residual connections on learning. *AAAI'17: Proceedings of the Thirty-First AAAI Conference on Artificial Intelligence*, 4278–4284. doi: 10.48550/arXiv.1602.07261

Thomas,, Fogel, D. B., and Michalewicz, Z. (1997). *Handbook of evolutionary computation*. (Boca Raton, Florida, USA: CRC Press). doi: 10.1201/9780367802486

Tran, L. T., and Bae, S.-H. (2021). Training-free hardware-aware neural architecture search with reinforcement learning. *J. Broadcast Eng.* 26, 855–861. doi: 10.5909/jbe.2020.26.7.855

Wang, Y., and Wang, Z. (2019). A survey of recent work on fine-grained image classification techniques. *J. Visual Commun. Image Representation* 59, 210–214. doi: 10.1016/j.jvcir.2018.12.049

Wang, B., Xue, B., and Zhang, M. (2020). "Particle swarm optimisation for evolving deep neural networks for image classification by evolving and stacking transferable blocks," in *IEEE Congress on Evolutionary Computation (CEC) as part of the IEEE World Congress on Computational Intelligence (IEEE WCCI) Electr Network*. pp. 1–8. doi: 10.1109/CEC48606.2020.9185541

Wen, Y.-W., Peng, S.-H., and Ting, C.-K. (2021). Two-stage evolutionary neural architecture search for transfer learning. *IEEE Trans. Evol. Comput.* 25, 928–940. doi: 10.1109/tevc.2021.3097937

Wistuba, M. AutoML (2017). *Automatic Machine Learning Workshop (ICML 2017)*. https://research.ibm.com/publications/bayesian-optimization-combined-with-incremental-evaluation-for-neural-network-architecture-optimization.

Xiao, X. Y., Hu, R., Zhang, S., and Wang, X. F. (2010). "Advanced Intelligent Computing Theories and Applications. With Aspects of Artificial Intelligence. ICIC 2010," In *Lecture Notes in Computer Science* 6216. Berlin, Heidelberg: Springer. doi: 10.1007/978-3-642-14932-0_19

Xie, L., and Yuille, A. (2017). "Genetic CNN," in *2017 IEEE International Conference on Computer Vision (ICCV)*. 2017, pp. 1388–1397. doi: 10.1109/ICCV.2017.154

Yu, X., Yankang, W., Jiayu, L., and Adam, S. (2021). A self-adaptive mutation neural architecture search algorithm based on blocks. *IEEE Comput. Intell. Magazine* 16 (3), 67–78. doi: 10.1109/MCI.2021.3084435

Zoph, B., and Le, Q. V. (2016) *Neural architecture search with reinforcement learning*. Available at: https://ui.adsabs.harvard.edu/abs/2016arXiv161101578Z.

frontiers | Frontiers in Plant Science

TYPE Original Research
PUBLISHED 30 November 2022
DOI 10.3389/fpls.2022.946154

OPEN ACCESS

EDITED BY
Lei Shu,
Nanjing Agricultural University, China

REVIEWED BY
Pavel Loskot,
The Zhejiang University-University of Illinois at Urbana-Champaign Institute, United States
Mohsen Yoosefzadeh Najafabadi,
University of Guelph, Canada

*CORRESPONDENCE
Yubin Lan
ylan@scau.edu.cn
Weixing Wang
weixing@scau.edu.cn

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

RECEIVED 17 May 2022
ACCEPTED 08 November 2022
PUBLISHED 30 November 2022

CITATION
Lu J, Yang R, Yu C, Lin J, Chen W, Wu H, Chen X, Lan Y and Wang W (2022) Citrus green fruit detection *via* improved feature network extraction. *Front. Plant Sci.* 13:946154. doi: 10.3389/fpls.2022.946154

Citrus green fruit detection *via* improved feature network extraction

Jianqiang Lu[1,2,3], Ruifan Yang[1,3], Chaoran Yu[4,5], Jiahan Lin[1,3], Wadi Chen[1,3], Haiwei Wu[1,3], Xin Chen[1,3], Yubin Lan[1,2,3]* and Weixing Wang[1,6]*

[1]College of Electronic Engineering (College of Artificial Intelligence), South China Agricultural University, Guangzhou, China, [2]Guangdong Laboratory for Lingnan Modern Agriculture, Guangzhou, China, [3]National Center for International Collaboration Research on Precision Agricultural Aviation Pesticide Spraying Technology, Guangzhou, China, [4]Vegetable Research Institute, Guangdong Academy of Agricultural Sciences, Guangzhou, China, [5]Guangdong Key Laboratory for New Technology Research of Vegetables, Guangzhou, China, [6]Guangdong Provincial Agricultural Information Monitoring Engineering Technology Research Center, Guangzhou, China

Introduction: It is crucial to accurately determine the green fruit stage of citrus and formulate detailed fruit conservation and flower thinning plans to increase the yield of citrus. However, the color of citrus green fruits is similar to the background, which results in poor segmentation accuracy. At present, when deep learning and other technologies are applied in agriculture for crop yield estimation and picking tasks, the accuracy of recognition reaches 88%, and the area enclosed by the PR curve and the coordinate axis reaches 0.95, which basically meets the application requirements.To solve these problems, this study proposes a citrus green fruit detection method that is based on improved Mask-RCNN (Mask–Region Convolutional Neural Network) feature network extraction.

Methods: First, the backbone networks are able to integrate low, medium and high level features and then perform end-to-end classification. They have excellent feature extraction capability for image classification tasks. Deep and shallow feature fusion is used to fuse the ResNet(Residual network) in the Mask-RCNN network. This strategy involves assembling multiple identical backbones using composite connections between adjacent backbones to form a more powerful backbone. This is helpful for increasing the amount of feature information that is extracted at each stage in the backbone network. Second, in neural networks, the feature map contains the feature information of the image, and the number of channels is positively related to the number of feature maps. The more channels, the more convolutional layers are needed, and the more computation is required, so a combined connection block is introduced to reduce the number of channels and improve the model accuracy. To test the method, a visual image dataset of citrus green fruits is collected and established through multisource channels such as handheld camera shooting and cloud platform acquisition. The performance of the improved citrus green fruit detection technology is compared with those of other detection methods on our dataset.

Results: The results show that compared with Mask-RCNN model, the average detection accuracy of the improved Mask-RCNN model is 95.36%, increased by 1.42%, and the area surrounded by precision-recall curve and coordinate axis is 0.9673, increased by 0.3%.

Discussion: This research is meaningful for reducing the effect of the image background on the detection accuracy and can provide a constructive reference for the intelligent production of citrus.

KEYWORDS

instance segmentation, Mask-RCNN, feature fusion, CB-Net, deep learning

1 Introduction

Citrus is an important cash crop in China, with an annual production of nearly 50 million tons. Scientific planning of fruit preservation and thinning is an important measure for ensuring citrus yield. During the green fruit stage of citrus, fruit development is easily affected by the environment, pests, and diseases, which results in deformed fruit, fruit with pests and diseases, and fruit with mechanical damage. The edible value of these fruits is very low, and there is little economic benefit. However, they absorb some of the nutrients of the fruit tree during the development process, which results in a waste of nutrients such that the normal fruit cannot obtain enough nutrient supply. At the same time, there are too many fruits on adult citrus trees, and the phenomenon of nutrient competition among fruits is serious. Therefore, it is important to accurately define the green fruit stage of citrus through scientific methods and to reasonably thin the fruit to improve the yield of citrus (Yan et al., 2021). The citrus green fruit stage is traditionally judged by the fruit grower's visual observation, which not only results in subjective judgment errors (Linker et al., 2012; He et al., 2016) but also has limitations for the intelligent and unmanned operation of orchard fruit thinning. With the development of smart agriculture applications, modern computer science and technology provide new strategies for crop target identification and detection, and real-time processing of orchard image data through sensor systems and high-performance computers (Yamamoto et al., 2014; Haseeb et al., 2020) can greatly reduce labor costs and improve detection accuracies. Therefore, it is important for growers to make orchard patrol plans according to the growth of citrus green fruit, and analyzes the pictures of citrus orchard taken by the camera using deep learning algorithm, so as to obtain the current number of citrus green fruits, so that growers can determine the yield of fruit trees and carry out timely fruit thinning operation.

Machine learning is used to accomplish the task of classification. Through supervised learning, fitting of a model to data (or a subset of data) that have been labelled– where there exists some ground truth property, which is usually experimentally measured or assigned by humans (Greener et al., 2022). Subsequently, this model is used to map all the inputs into the corresponding outputs and make a simple judgment on the outputs for prediction and classification, which also has the ability to predict andclassify the unknown data. Yoosefzadeh et al. (Yoosefzadeh Najafabadi, 2021) implemented ML algorithms in GWAS, investigated the potential use of RF and SVM algorithms in GWAS to detect the associated QTL with soybean yield components, which would be beneficial to select the superior soybean genotypes. Therefore, integrating artificial intelligence and computer vision technology to establish a citrus green fruit stage detection model would be an effective smart agriculture approach (Yang et al., 2021) for detecting citrus (Rakun et al., 2011; Lin et al., 2019).

For target detection and recognition of citrus fruits, traditional machine learning methods mainly use learning algorithms such as edge detection algorithms, watershed segmentation algorithms, and support vector machine algorithms (Kurtulmus et al., 2011; Lu and Hu, 2017; Peng et al., 2021). Such methods design feature extraction algorithms for the color, texture, and shape of crops or agricultural products (Li et al., 2017), segment relevant features in steps, and accurately locate targets in images (Wang L et al., 2022). For example, (Dorj et al., 2017) proposed a color feature-based citrus yield estimation algorithm.

Abbreviations: Mask-RCNN, Regions with Convolutional Neural Network features; Res4b, A module in the Residual network; ResNet, Residual network; CB-Net, Composite Backbone Network; COCO, Microsoft Common Objects in Context; HLB, HuangLongBing disease; RPN, Region Proposal Network; Faster-RCNN, Faster Region Convolutional Neural Networks; VGG16, Visual Geometry Group Network16; CNN, Convolutional Neural Network; FPN, Feature Pyramid Network; RoI, Region of Interest; CC, composite connection; AHLC, Adjacent high-level composition; AP, Average Precision; IoU, Intersection over Union.

Based on the automatic watershed algorithm, distance conversion and marker control methods have been introduced, which can better segment the individual citrus fruits in images. (Hu, 2018) improved machine learning-based citrus green fruit detection by using the local binary method and the maximum stable polar region algorithm to extract the color images in the region of interest, using the Hough transform to fit each level of contour lines to obtain a hierarchical circular target, and finally, performing a fitted circle nested analysis to obtain the citrus green fruit target. The above methods describe individual features of field crops in color, texture, and shape space to achieve target and background segmentation.

These traditional field crop detection methods require high background complexity of the input image, and their performance on complex and diverse agricultural orchard scenes is limited. In addition, these methods perform feature extraction on a single scale and fail to produce high-accuracy detection results. Instance segmentation algorithms provide pixel-level target detection methods for solving the problem of inaccurate classification due to individual deformation of targets of the same category in target detection methods and achieving the detection of different individuals of targets of the same category (Liu et al., 2018; Wang et al., 2020; Jia et al., 2021). The most widely used instance segmentation algorithm is the Mask-RCNN (He et al., 2017) algorithm (Wang et al., 2016), which applies the extended convolution method to the Res4b module of ResNet, which is the backbone network of Mask-RCNN, for the recognition and localization of poplar plum in the natural environment to achieve accurate recognition and segmentation of poplar plum. Zhang Y et al. (Zhang, 2020) used the Mask-RCNN algorithm with a Kinect V2 (Lachat et al., 2015) camera to acquire apple images under different environmental conditions and to segment the generated apple point cloud data (Wahabzada et al., 2015). Deng Y et al. (Deng et al., 2020) achieved efficient detection of dense small-scale citrus flower targets in complex structured images and acquired the number of visible flowers in images by optimizing the body convolution part and the mask branching part of the Mask-RCNN algorithm. The Mask-RCNN algorithm has high-efficiency detection performance and high operability and is widely used in various field crop detection and segmentation (Santos et al., 2020) tasks. It is an important tool for implementing instance segmentation tasks in agriculture.

Current research has focused on target recognition of ripe citrus yellow fruits, and less research has been conducted on the detection of citrus green fruits (Zhao et al., 2016), but it is important to accurately identify citrus green fruits and define the citrus green fruit stage. Compared with ripe citrus yellow fruits, citrus green fruits are more difficult to recognize. The reasons are as follows: (1) citrus green fruits are difficult to distinguish because their color is similar to the background under natural light (2) under natural conditions, citrus green fruits are small in size and occupy very few pixels in the image (3) citrus green fruits overlap each other and thus are easily blocked by leaves, branches and other background objects, which is difficult to detect. Therefore, we need to more accurately extract the inherent characteristics of citrus green fruits, fuse the multiclass features of citrus green fruit, and use efficient feature extraction methods to improve the accuracy of citrusgreen fruit detection and segmentation.

Deep convolutional neural networks are constantly evolving, and many backbone networks are able to integrate low, medium and high level features and then perform end-to-end classification. They have excellent feature extraction capability for image classification tasks, and common backbone networks are VGG, Resnet, etc. In neural networks, each channel needs to do convolution operation with a convolution kernel, and then the results are summed to get a feature map output. The feature map contains the feature information of the image, and the number of channels is positively related to the number of feature maps. The more channels, the more convolutional layers are needed, and the more computation is required, so reducing the number of channels is beneficial to the computation speed.

This study selects citrus green fruits in the natural environment as the research objects due to the limited accuracy of traditional target detection algorithms for detecting citrus green fruits in complex backgrounds. Based on an improved version of the pixel-level instance segmentation algorithm Mask-RCNN (Fan et al., 2021), a citrus green fruit detection method is designed by introducing CB-Net (Composite Backbone Network) (Liu et al., 2020). The method involves assembling multiple identical backbones using composite connections between adjacent backbones to form a more powerful backbone. This helps increase the feature information that is extracted at each stage in the backbone network. Then, a combined connection block is introduced to reduce the number of channels and improve the model accuracy. This method can effectively mitigate the problem that the citrus green fruit color is similar to the background color, which reduces the detection accuracy. The proposed algorithm is pretrained by combining the data of citrus green fruit images that were captured from multiple angles using a camera and cloud platform with the training weight file of the Mask-RCNN algorithm on the COCO dataset. Then, the algorithm is formally trained and tested on the collected citrus green fruit images for evaluation. This study provides a research basis for the tasks of detecting and dividing citrus green fruits under natural conditions and the development of intelligent and unmanned operations for citrus green fruit thinning, and it broadens the research scope of intelligent agriculture in the field of citrus flower and fruit preservation. It is important to improve the efficiency of citrus operations and promote the development of citrus production.

In this study, the image data of citrus green fruit under real natural environment was collected, and the corresponding data enhancement processing was carried out to construct a citrus green fruit data set. Based on the Mask-RCNN algorithm, CB-Net with deep and shallow fusion was innovatively combined with the

traditional feature extraction network ResNet to fuse the multilayer features of citrus green fruits. Then, the proposed model was compared with the traditional model in detail. Theresults showed that the proposed algorithm based on the improved Mask-RCNN has improved citrus green fruit detection accuracy and speed.

The main contributions of this study are as follows:

(1). Based on the strategy of feature fusion, an instance segmentation method is proposed for reducing the influence between citrus green fruit features and irrelevant features.

(2). We construct a novel Mask-RCNN model using CB-Net (a composite backbone network) to fuse the multilayer features of citrus green fruits so that the model can focus more on the obvious regions and more detailed features in each image.

(3). The performance of the proposed model in detecting different individuals under the same class of target to obtain the morphology of citrus green fruits in advance is evaluated.

2 Materials and methods

2.1 Study area

Sugar tangerine is an important variety in China's citrus industry. Its main production area is South China, and it has high edible value. Its growth stage is similar to that of ordinary citrus. The cultivation of sugar tangerine has high environmental requirements, and sugar tangerine has poor resistance to insects and diseases. It is necessary to observe the growth of each tree for a long time to supplement nutrients in time and promote the normal development of flowers, fruits, and leaves (Rahnemoonfar and Sheppard, 2017). Using computer vision to detect citrus green fruit targets requires high-resolution images and imaging data to study the important shape features of citrus green fruit and to finally evaluate the accuracy of the results. The experimental site for this study is located in the research and development demonstration base for Green Plant Protection of Citrus HLB (Huanglongbing) and New Cultivation Modes in Jingshuilong Village, Yangcun Town, Boluo County, Huizhou City, Guangdong Province (N23° 29'57.81"—N23°29'59.31",E114°28'8.39"—E114°28'12.26"). It is 40 m above sea level, and the local climate is mild and humid, which is suitable for the planting of citrus and other fruit trees. The crop varieties in this test area are all sugar tangerines. There are 334 citrus plants in the test area, with 9 rows, a row spacing of 4 m, and a column spacing of 2.5 m. In the natural light environment, visible image data of citrus green fruits were collected in July 2020. The citrus experimental site is shown in Figure 1.

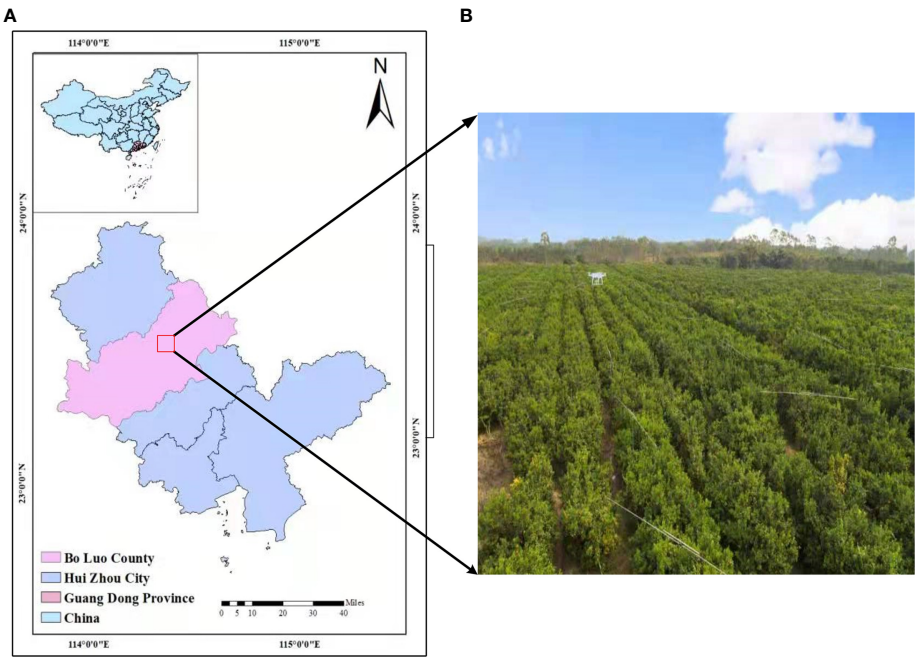


FIGURE 1
Study Area: **(A)** Geographical location of the study area and **(B)** the citrus test base.

2.2 Test data

2.2.1 Test data collection

To improve the quality and diversity of citrus green fruit images, the modes of the collection were as follows: 1. A manual hand-held camera (Model D7100, APS-C frame camera of Nikon, Japan) was used to take multiangle visible images of citrus green fruits at a distance of 2-3 m from the canopy of the citrus tree, which produced JPEG images with a resolution of 4928×3264; 2. A wireless zoom camera (Hikvision 3T27EWD) was called in the orchard through the cloud platform. This zoom camera has a 1/2.7 largetarget sensor. At a distance of 5-10 m from the canopy of the citrus tree, we used it for remote real-time acquisition of visible images of citrus green fruits, which produced JPEG images with a resolution of 1280×720. The data acquisition mode diagram is shown in Figure 2. Finally, 200 and 357 citrus green fruit images were collected using the hand-held camera and cloud platform, respectively, for a total of 557 images. The number of citrus green fruits in each image was 1-10. To obtain a unified data format, the picture resolution was compressed to 1280×720. Figure 3 shows citrus green fruit picture data that were collected by the two data collection modes.

2.2.2 Data processing

The data in supervised learning needs to be classified in advance, and its training samples contain both feature and label information. Therefore, a method based on improved Mask-RCNN for constructing citrus green fruit dataset is proposed. First, Labelme (Russell et al., 2008) data labeling software is used for instance labeling of citrus green fruit individuals, and the interface of the data labeling software is shown in Figure 4. The top of the software is the menu bar, the left side is thetoolbar, including open file or folder, select the current picture before and after the picture file, save the annotation file, select the annotation method, etc., a picture annotation area in the middle, and the right side is the picture and the annotation name, category and other information. When labeling, select polygon labeling method to manually mark the citrus target with dense dotting, and mark the visible citrus target area.

The citrus green fruit of each monomer is regarded as a class. When there are multiple citrus targets in the Figure 4, the labels are set to “cirtus1”, “cirtus2”, “cirtus3”, and so on. When the labeling of an image is completed, a json label filewith the suffix “.json” is generated, which records the version number of Labelme software, the label of each citrus green fruit labeled and the pixel coordinates of the corresponding labeling point.

In this study, we fully consider the various shapes in the environment in which the citrus green fruits are located in the sample annotation process and ignore the citrus targets that are obscured by more than 70%. The method can accurately obtain the citrus green fruit target locations and reduce the interference of citrus green fruit with obscured feature information in images with complex backgrounds. A total of 3273 Citrus green fruit samples are labeled.

In this study, data augmentation is used to improve the network learning and generalization ability of the network model. We mainly use image rotation, image horizontal flipping, vertical flipping, and horizontal-vertical flipping as data augmentation methods. Rotating and flipping images can improve detection performance. Meanwhile, hybrid augmentation is designed to address the limitation of the overdependence of the model on the dataset. The mixture of different classes of samples in the dataset is used to generate new samples, which enhances the linear expression among different classes of samples and improves the robustness against the samples. The size of the amplified dataset is 2228. Furthermore, the dataset is randomly divided into a training set, validation set, and test set at a ratio of 6:2:2 for training, tuning, and testing. A flow chart of the data preprocessing is shown in Figure 5.

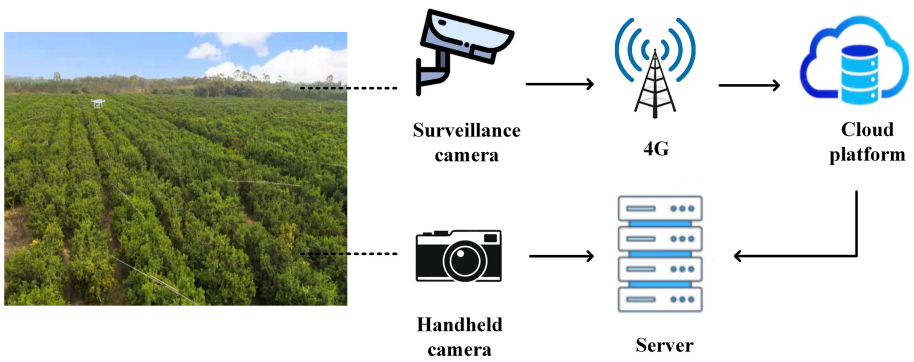


FIGURE 2
Data acquisition mode diagram.

FIGURE 3
Citrus green fruit images. **(A)** Pictures of citrus green fruit that were captured with a manual handheld camera and **(B)** pictures of citrus green fruit that were obtained by the cloud platform.

FIGURE 4
Data annotation interface.

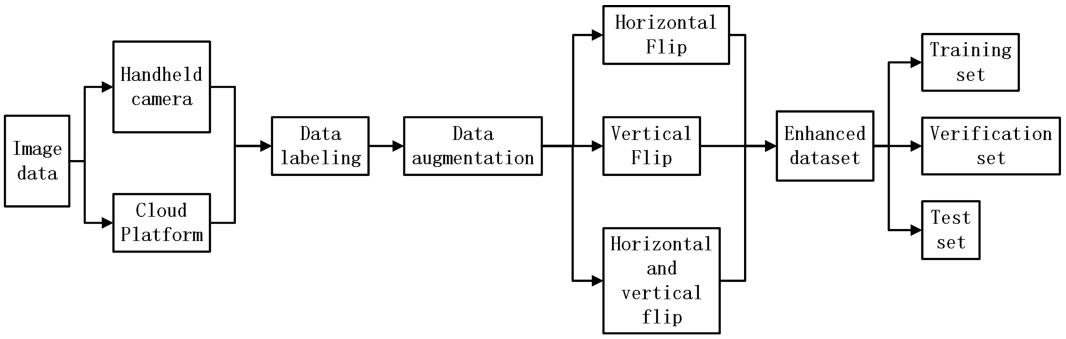


FIGURE 5
Data pre-processing flow chart.

2.3 Mask-RCNN

In deep learning, the semantic information includes the texture, color, or category of the target in the image, the richer the semantic information, the stronger the correlation between each pixel point and the surrounding pixels in the image. In the process of citrus green fruit recognition, it is necessary to preserve and integrate the feature maps with different resolutions that are generated by the feature extraction network in each convolution stage to generate feature maps with rich high-resolution semantic information.

This is especially beneficial for improving the recognition rate of green fruits and distinguishing the subtle features of fruits and leaves. In the current research, the RPN structure that is proposed by Faster-RCNN (Ren et al., 2015; Sa et al., 2016; Apolo-Apolo et al., 2020) has advantages in terms of model accuracy and training prediction speed, but traditional unidirectional networks such as ZFNet (Fu et al., 2018) and VGG16 (Qassim et al., 2018) are adopted by the CNN feature extraction network to convolve and sample the original images with high resolution and weak semantic information from top to bottom and finally generate feature maps with low resolution and strong semantic information. When the target shape or feature difference in the image is small, a low-resolution feature map will easily lose the feature information of small targets, thereby resulting in a decreased recognition rate and missed detection of the small target, among other effects.

The Mask-RCNN model is based on Faster-RCNN with the addition of a semantic segmentation branch for outputting the mask of the target and adjusting the training parameters through the loss function to achieve deep learning of image features. The Mask-RCNNalgorithm introduces a feature pyramid network into the CNN feature extraction network (Lin et al., 2017) and uses a ResNet network that is based on residual learning as the feature extraction network. In contrast, FPN in the Mask-RCNN algorithm fuses multiple feature scales and semantic information, which can realize multiscale feature extraction and fusion of images.

Mask-RCNN adds a mask prediction branch to the target detection algorithm Faster-RCNN and performs convolution and fully connected operations on the feature map in parallel with the bounding-box regression branch and classification branch. Moreover, it uses the RoI Align (Wang and He, 2019) method instead of RoI Pooling (Liu et al., 2017) of Faster-RCNN to enhance the pixel-to-pixel correspondence between network inputs and outputs, reduce the error of the bounding-box regression, and improve the target detection accuracy (Li et al., 2020). Based on the above features, Mask-RCNN enhances the feature information between citrus green fruit and the background in the process of detecting citrus green fruit, which is helpful for reducing the difficulty of citrus green fruit detection and segmentation. Therefore, this study explores the high-resolution optimization of feature maps in citrus green fruit detection based on the advantages of Mask-RCNN, which fuses multiple feature scales and semantic information to achieve multiscale feature extraction and fusion of images (He et al., 2020).

2.4 The proposed algorithm

The feature extraction network extracts the shape features of citrus green fruit by a convolution operation and builds a multilayer neural network model to realize the recognition and localization of citrus green fruit in images. However, in the actual scene, the citrus green fruit and the leaves are similar in color, and some of the leaves are also round-like in outline, which makes it difficult for the model that is built by a single feature extraction network to distinguish the feature information of the citrus green fruit and background, which increases the difficulty of detection and segmentation of citrus green fruit. Instance segmentation has both the characteristics of semantic

segmentation and target detection. The region where the instances are located is identified by the target detection method, and then semantic segmentation is performed within the detection frame, and each segmentation result is output as a different instance. Since citrus green fruits differ from leaves in shape and color by subtle features, it is necessary to design a deep and shallow feature extraction network with both extraction and fusion functions in order to describe the inherent features of citrus green fruits more accurately. The improved algorithm is used to further extract phenotypic features such as shape and color of citrus green fruits under a green background. At the same time, the object detection is further refined to fuse the multiple classes of features of citrus green fruits. Then, the extracted multi-scale feature information is used to separate the detection object from the background and achieve accurate segmentation at the pixel level.

In this study, based on the Mask-RCNN network structure, CB-Net is introduced. CB-Net provides a highly effective feature extraction method for target detection and instance segmentation algorithms based on the strategy of composite connection, which is a worthwhile optimization strategy for tasks in which detection is difficult and the feature effect is not obvious. In this study, we expect the improved algorithm to effectively identify citrus green fruits in similar background environments and obtain better model accuracy at the expense of the time cost of model training and prediction.The overall structure of the Mask-RCNN model that incorporates the CB-Net strategy is illustrated in Figure 6. The backbone network consists of a feature extraction network that uses the ResNet + CB-Net network and backbone, RPN, and an ROI head. The input citrus green fruit images are compressed and passed into three branches of the improved feature extraction network. ResNet based on CB-Net is used to extract and fuse multiple phenotypes of citrus green fruit, and the expression ability of multiscale features is enhanced by FPN. Furthermore, the feature maps are generated and corrected for candidates bound by the RPN and ROI align modules. Finally, three prediction branches of regression, classification (Li et al., 2019) and masking are used for the detection and segmentation of citrus green fruit. The joint loss function of Mask-RCNN is used to optimize the parameters in the model training process.

The improved Mask-RCNN uses the ResNet + CB-Net network as the feature extraction network, and the structure of the network is illustrated in Figure 7. The design strategy is that ResNet50 or ResNet101 (Hong et al., 2020) is iterated many times,and a composite connection module is used between each ResNet block for transverse propagation of the feature maps, which can effectively increase the amount of feature information that is extracted at each stage in the backbone network and improve the performance of citrus green fruit detection in similar background environments.

The stage names of each ResNet network are Ci-j in Figure 7, where i denotes the i-th ResNet network and j denotes the j-th stage. The CC module is a composite connection module, and the numbers of channels for 1×1 convolution in CC1, CC2, CC3, and CC4 are 64, 256, 512, and 1024, respectively.

A single ResNet network has fixed requirements for the size of the input images. Before images are input to the ResNet network, they need to be resized. In this study, the original images of citrus green fruit are resized to 1024×1024×3 as the first-stage input of each ResNet network. As the depth of the network increases, the size of the feature maps decreases, and the output size of stage C1 is 256×256×64, that of stage C2 is 128×128×256, that of stage C3 is 64×64×512, that of stage C4 is 32×32×1024, and that of stage C5 is 16×16×2048. An important parameter of ResNet + CB-Net is the number of network iterations. With the increase in the number of network iterations, the final extracted feature information becomes

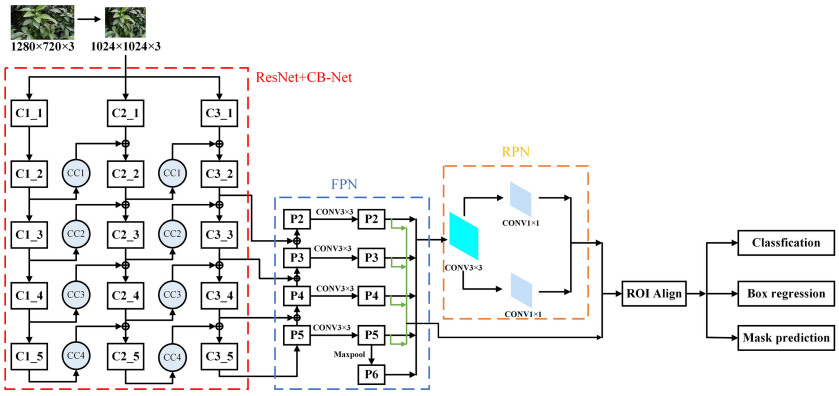


FIGURE 6
The structure of Mask-RCNN model integrating CB-Net idea.

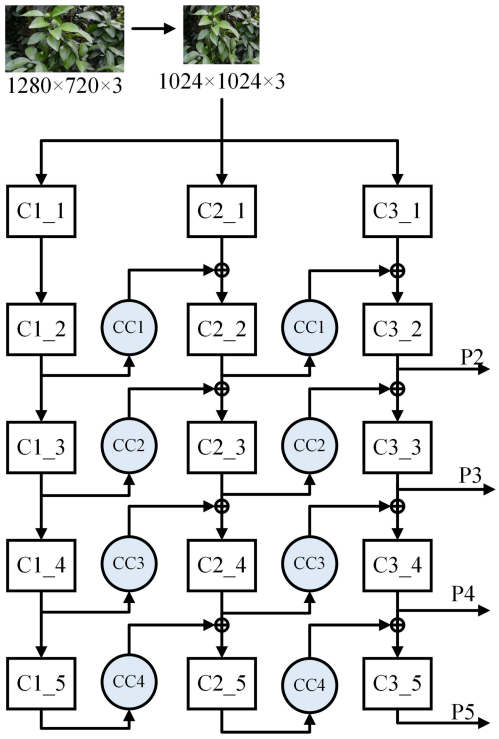


FIGURE 7
ResNet+CB-Net Network structure.

richer, and the model expression ability improves gradually. When the number of iterations reaches a threshold value, the improvement in the network accuracy decreases gradually, and as the number of iterations is further increased, the network accuracy improvement becomes close to zero. The number of iterations is linearly related to the time of model training and the capacity of the physical memory that is occupied by the hardware. When the accuracy of the network reaches the saturation condition, increasing the number of iterations has little benefit in terms of the accuracy but greatly reduces the operational efficiency of the model and increases the ratio of the storage space that is occupied by the model. To ensure that the entire network is highly efficient, the number of iterations is set to 3; that is, three identical ResNet networks are used for connections.

The network adopts AHLC mode for composite connections. The composite connection module consists of a convolutional layer with a convolutional kernel of size 1×1, a batch normalized layer (Ioffe and Szegedy, 2015), and an upsampling module. The 1×1 convolutional operation changes the number of channels. Batch normalization can improve the model efficiency and reduce regularization processing. The upsampling module changes the size of the feature image for linear matrix superposition with the convolutional layer of the next ResNet network. In one connection, we assume that the output of stage i of the first ResNet network is O1 (i). After a composite connection with C(·), O1 (i) is superimposed with the output of stage i-1 of the second ResNet network as the input I2 (i) of Phase i of the second ResNet network. Then, the output O2 (i) of Phase i of the second ResNet can be obtained by Formula 1, where F(·) is a convolutional operation of stage i.

$$O_2(i) = F(I_2(i)) = F(C(O_1(i))) + O_2(i-1), i \geq 2 \qquad (1)$$

To further demonstrate the process of lateral transmission of feature information, the output of C1-2 is transmitted to the second ResNet network through a composite connection module as an example (Figure 8). C1-2 is the second stage of the first ResNet network, so its output feature map size is 128×128×256. The number of channels of the feature map is downsampled by CC1 with a 1×1×64 convolution, and the size of the output feature map is 128×128×64. In addition, to perform linear summation with the output feature map of C2-1, the length and width parameters of the feature map are further upsampled. Then, the size of the output feature map of CC1 is 256×256×64, which is the same as the size of the output feature map of C2-1. To generate multiple outputs from the backbone, a composite connection module is introduced. This module consists of a 1x1 convolutional layer and a batch normalization layer. These

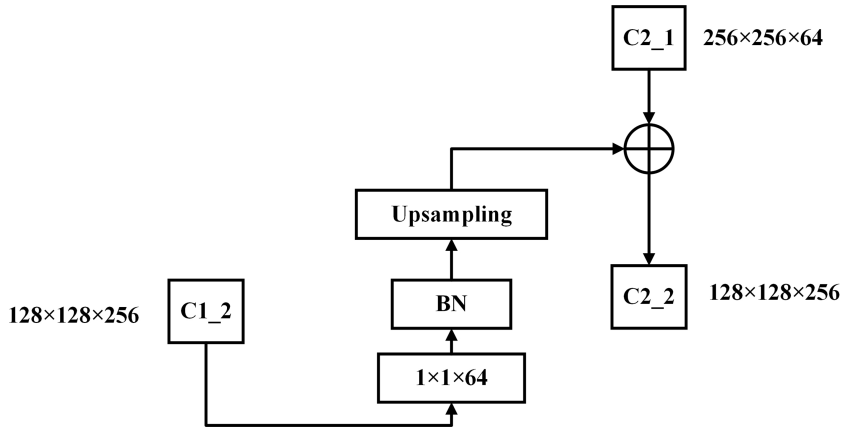


FIGURE 8
Composite connection module.

layers are added to reduce the number of channels and to perform an upsampling operation to increase the feature image resolution and improve the model accuracy. To further demonstrate the process of lateral transmission of feature information, the output of C1-2 is transmitted to the second ResNet network through a composite connection module as an example (Figure 8) C1-2 is the second stage of the first ResNet network, so its output feature map size is 128×128×256. The number of channels of the feature map is downsampled by CC1 with a 1×1×64 convolution, and the size of theoutput feature map is 128×128×64. In addition, to perform linear summation with the output feature map of C2-1, the length and width parameters of the feature map are further upsampled. Then, the size of the output feature map of CC1 is 256×256×64, which is the same as the size of the output feature map of C2-1.

To obtain the RPN structure and the feature maps that are required for the ROI prediction branches, P2~P5 of the output of the last ResNet network are input into the FPN.

2.5 Performance evaluation metrics

In this study, precision rate (P), recall rate (R), and precision-recall curve (PR) are used as evaluation metrics. Precision refers to the prediction result, which is defined as the number of targets that are true positive examples among the targets predicted as positive examples. Recall refers to the sample data, which is defined as the number of targets that are predicted as positive examples among the targets predicted as positive samples. In equations 2 and 3, TP denotes the number of positive samples with correct predictions, FP denotes the number of positive samples with incorrect predictions, and FN denotes the number of negative samples with incorrect predictions. The PR curves allow us to analyze the dynamic trends of the accuracy rate and recall rate on the whole data set and compare the performance of different models on the same data set. The calculation formulas are as follows:

$$P = \frac{TP}{TP + FP} \tag{2}$$

$$R = \frac{TP}{TP + FN} \tag{3}$$

AP summarizes the shape of PR curve from the numerical level, and its value is the average of accuracy at the recall level with equal intervals of 0 to 1. The calculation formula is:

$$AP(c) = \int_0^1 Precision\ \ rate(c)\ \ dRecall\ \ rate(c) \tag{4}$$

where C represents the goal of a category. In order to evaluate the effectiveness of the proposed Citrus green fruit detection method, the precision-recall curve and average accuracy AP50 and AP75 are used as evaluation indexes. AP50 and AP75 represent the average accuracy when IoU threshold is set as 50 and 75.

2.6 Model training and testing

The computer hardware configuration parameters that are used in this test are as follows: The operating system is Ubuntu18.04, the processor is an Intel Xeon(R) CPU E5-2620 V4 @ 2.1 GHz ×16, and the memory is 64 GB. The graphics processing unit (GPU) is a GTX TIAN X.

Migration learning (Shilei et al., 2019) is a machine learning method for knowledge domain migration. The core strategy is as follows: Knowledge models of mature domains, which are obtained by algorithms learning on massive data for a long

time, are applied to the model training of new domains, where the similarity between data and domains is used to share parameters with models of new domains to reduce hardware resource consumption, while migration learning builds rich low-level semantic featuresfor the models and guides the models to learn appropriate weighting parameters parameters (Yin et al., 2020). The citrus green fruit dataset that is collected in this study is obviously insufficient to support the data size that is required for learning the parameters from no initial information.

Based on the above strategy, this study utilizes the pretraining weights of the Mask-RCNN algorithm on the COCO dataset to pretrain the model in the initial stage of model training and then fine-tunes the model using the citrus green fruit dataset that isestablished in this study (Wang and Xiao, 2021). The training is backpropagated using minibatch gradient descent (MBGD) to optimally update the model parameters. During training, the batch size is set to 2, and the network model initialization learning rate is set to 0.001. It lasts for 30 epochs in total. After training, the loss values of the model are recorded after each iteration, and the correlation curves between the number of iterations and the loss values are plotted to analyze the accuracy change and convergence of the model during training. The final converged model is saved, and then model prediction is performed on the test set. The average precision, accuracy, and recall are calculated and saved, and the PR curve is plotted to analyze the generalization ability of the model.

3 Results

3.1 Performance comparison of different algorithms

To evaluate the effectiveness of the data enhancement method in solving the overfitting problem, the Mask-RCNN algorithm is used for training and testing on the citrus green fruit images before and after data enhancement. After training on 557 original images, the model has an AP value of 71.31% on 446 test set images, and after training on 1782 images with data enhancement, the model has an AP value of 92.47% on 446 test set images, as presented in Table 1.

Based on the data-enhanced citrus green fruit dataset, the Mask-RCNN algorithm integrated with CB-Net is compared with the traditional Mask-RCNN algorithm. ResNet50 and ResNet101 are selected as the feature extraction backbone networks. Four groups of experiments are conducted, namely, Mask-RCNN +ResNet50, Mask-RCNN+ResNet101, Improved Mask-RCNN +ResNet50, and improved Mask-RCNN+ResNet101.

TABLE 1 Data enhancement test results.

Models	AP
Mask-RCNN (original)	71.31
Mask-RCNN (enhanced)	**92.47**

AP: Average Precision.
Bold values represent the best indicators in the current table.

The loss value variance curves of the four groups of tests are shown in Figure 9. The loss values of all four models eventually converge between 0.4 and 0.5, and the four models reach better values within 30 training iterations.

3.2 Comparison of improved model with other models

To evaluate the performance of the proposed model, training is performed on the citrus green fruit dataset that is constructed in this study, and the improvement points are compared one by one. The results are presented in Table 2. The average accuracy of the improved Mask-RCNN+ResNet50 model on the citrus green fruit dataset that is built in this study is 95.36, which is 1.42%, 3.13%, and 2.17% better than those of the other three models. The results from the segmentation prediction are visualized in Figure 10. The number of real citrus green fruits in the original image is 5. The improved Mask-RCNN+ResNet50 model correctly and completely identifies all citrus green fruits. The other three models misjudge the green leaves as citrus green fruits or miss the detection of citrus green fruits. Among them, the Mask-RCNN+ResNet50 model misjudges 3 and misses 1, the Mask-RCNN+ResNet101 model misjudges 5, and the improved Mask-RCNN+ResNet101 model misses 1. The results show that the improvedMask-RCNN+ResNet50 model greatly improves the accuracy of the detection and segmentation of citrus green fruit.

The two-stage target detection algorithm usually first uses the algorithm (selective search or region proposal network, etc.) to extract candidate frames from the image, and then performs secondary correction on the candidate frame target to obtain the detection result. Therefore, the accuracy of candidate frame selection is particularly important for the target detection task. In the feature extraction stage, this study mainly uses the adoption of FPN to enhance the expression of multi-scale features of citrus green fruits. In a further step, the feature map is passed through the RPN and ROI Align modules to generate and correct candidate frames. In order to verify the accuracy of the candidate frames of the improved model, we calculate the accuracy and recall rates of the four groups of experiments on the test set and plot the PR curves, as shown in Figure 11. The area that is enclosed by each PR curve and the coordinate axes reflects the accuracy of the candidate frame. When the area is larger, the confidence of the candidate bound that is predicted by the model is higher. The areas that are enclosed by the PR curve and the coordinate axes of the four models are calculated, and the results are presented in Table 3. The area that is enclosed by the PR curve and the coordinate axes

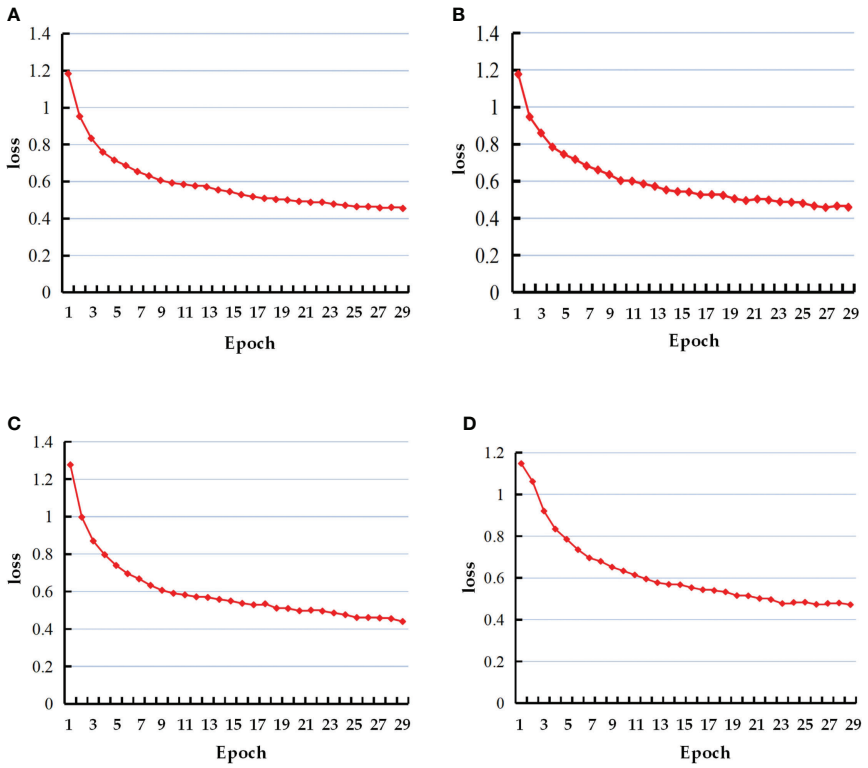


FIGURE 9
Loss curves of the four models. **(A)** Mask-RCNN+ResNet50; **(B)** Mask-RCNN+ResNet101; **(C)** Improved Mask-RCNN+ResNet50; **(D)** Improved Mask-RCNN+ResNet50.

of the improved Mask-RCNN+ResNet50 model is 0.9673, which is 0.3%, 1.3%, and 1.2% larger than the areas of the other three models.

4 Application case study

Citrus green fruit thinning assisted identification detection was conducted in the citrus orchard using the improved model that was suggested in this paper and transplanted to NVIDIA Agx Xavier. The citrus orchard is located in Mutan Village, Zengcheng City, Guangdong Province, China (as shown in Figure 12A). The specific parameters of the equipment used are as follows: NVIDIA Agx Xavier, the processor model is 8-core ARM v8.2, the memory model is eMMC5.1, the size is 32GB, and the GPU (Graphics Processing Unit) is 512-core Volta GPU with Tensor Cores. High-definition camera (Daipu DP-UK100), the size of HDCMOS sensor is 1/2.8 inches, the effective pixel is 2.1 million, and it can shoot 1080P ultra-high-resolution images.

TABLE 2 Prediction results of four sets of trials.

Models	AP_{50}	AP_{75}
Mask-RCNN+ResNet50	94.03	92.17
Mask-RCNN+ResNet101	92.47	89.87
Improved Mask-RCNN+ResNet50	**95.36**	**93.45**
Improved Mask-RCNN+ResNet101	93.34	91.68

AP_{50}: The IOU threshold is 0.5.
AP_{75}: The IOU threshold is 0.75.
Bold values represent the best indicators in the current table.

Before deploying the model with an edge smart station, it is necessary to initialize the edge smart station. In order to facilitate subsequent debugging and development, the Jetpack component library is used to reinstall the Agx Xavier system. The reinstalled operating system is Ubuntu 18.04, with GPU acceleration application tools such as CUDA 10.2.89 and Cudnn 8.0.0.180, and visual computing tools such as OpenCV and Vision Works. On this basis, according to the CUDA version, install thedeep learning environment required by Mask RCNN. The environment is as follows: Tensorflow gpu 1.15.0, keras 2.1.3, scikit image. Then, we connect the high-definition camera to the edge smart station through the USB interface, and configure the camera interface. We use the OpenCV toolkit to obtain the real-time image data of the camera, and call the model to detect each image acquired in real time.

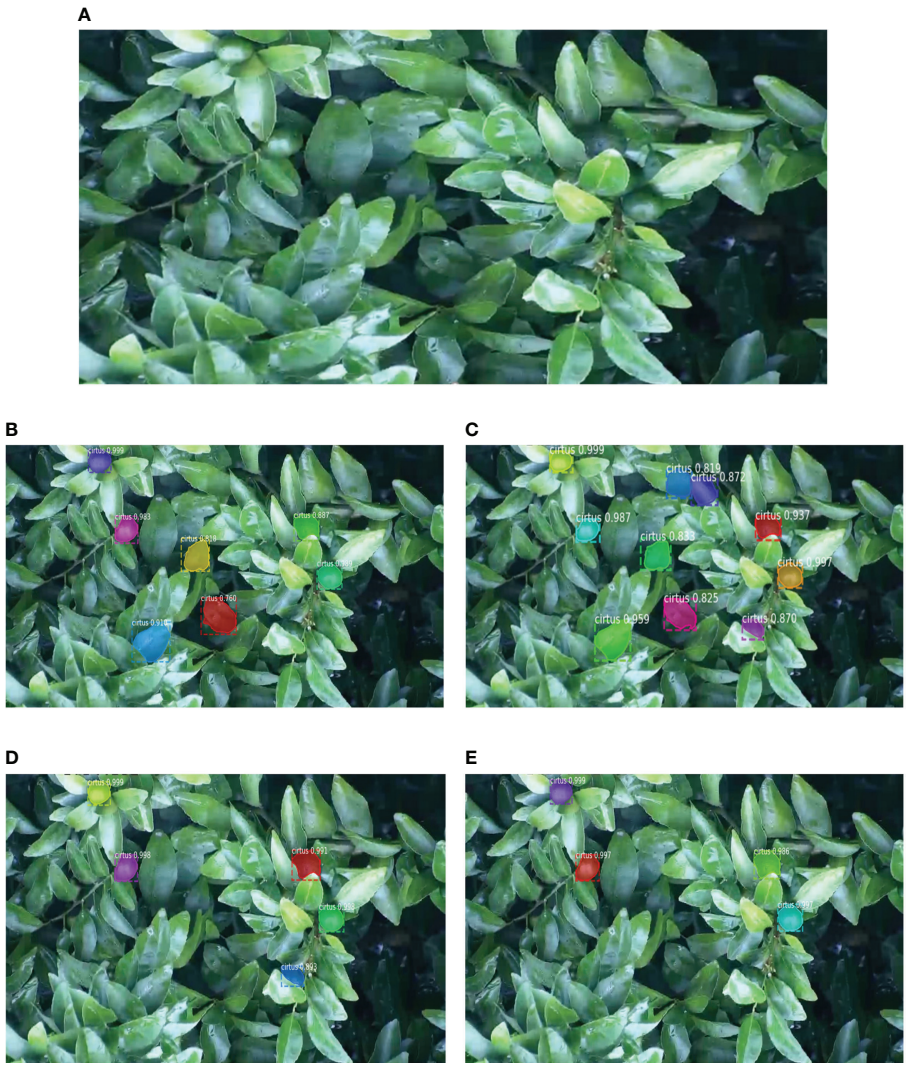


FIGURE 10
Forecast visualization results of the four models. **(A)** Original image; **(B)** Mask-RCNN+ResNet50 model; **(C)** Mask-RCNN+ResNet101 model; **(D)** Improved Mask-RCNN+ResNet50 model; **(E)** Improved Mask-RCNN+ResNet101 model.

Place Agx Xavier and the camera on a tripod at a distance of 2-3 meters from the citrus canopy and 1.6 meters above the ground (as shown in Figure 12B). A total of 10 collection points were set up in the orchard, and the focal length and angle of the camera were adjusted according to the position of the equipment and the green fruit. The camera took 2 hours each time. Both data collection and model reasoning were processed on Agx Xavier, and the processed results would be analyzed by the systemto further give fruit thinning decision-making suggestions. The overall flow chart is shown in Figure 12C. After the improved model was transplanted to Agx Xavier, the recognition of green fruits was in line with the performance indexes achieved during the model training test, and the inference time was 0.939 seconds, which could satisfy the demand for automatic green fruit recognition in citrus orchard thinning operations.

The equipment need to be manually moved about to collect various tree plants because it is mounted on a fixed tripod with a small field of view. In order to expand the applicability of the system in the future and further increase the robustness and feasibility of the system, the improved model proposed in this paper can be deployed on an unmanned vehicle in the future (as shown in Figure 12D). The method is as follows, an NVIDIA Agx Xavier that has been ported with an improved model and a HD camera is built on the unmanned vehicle. Then, by manually setting fixed trajectories in citrus orchards in advance, the unmanned vehicle will move along the preset trajectories when performing the task of citrus green fruit yield detection in citrus

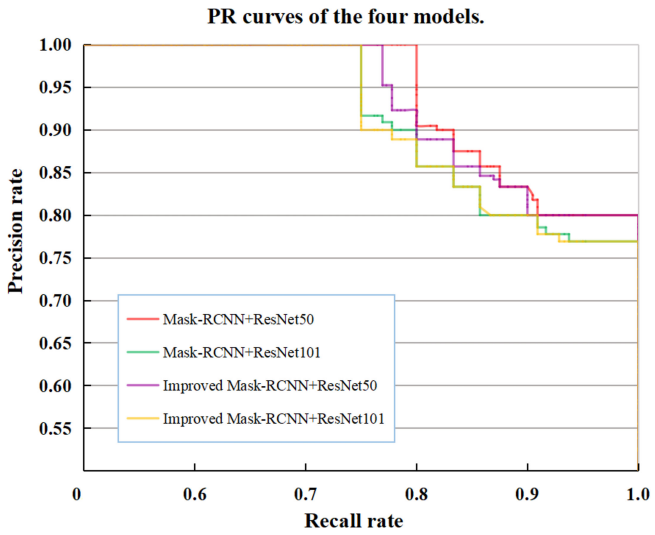


FIGURE 11
PR curves of the four models.

orchards. During the task, the following operations are repeated: (1) the unmanned vehicle is stationary once every 2 minutes (2) the HD camera takes pictures to obtain real-time image data (3) Agx Xavier invokes the model to detect each image obtained in real time. This enables the acquisition of image data of the whole orchard with low labor cost.

5 Discussion

To address the problem of difficult identification of citrus green fruits in the natural environment, an instance segmentation method based on the strategy of feature fusion is proposed for reducing the influence between citrus green fruit features and irrelevant features, and we construct a novel Mask-RCNN model using CB-Net (a composite backbone network). The generalization ability of the neural network was improved by constructing the green citrus dataset and a series of dataset preprocessing operations. Second, a combined connection block is introduced to reduce the number of channels and improve the model accuracy. The accuracy of the final trained model on the test set is 95.36%, which is higher than that of Mask-RCNN. This indicates that the improved model can more accurately identify more citrus green fruits.

To further verify the effectiveness and feasibility of this method, the performance of the proposed model in detecting different individuals under the same class of target to obtain the morphology of citrus green fruits in advance is evaluated. Compared with the original method, the average detection accuracy of the improved Mask-RCNN model is 95.36%, increased by 1.42%, and the area surrounded by precision-recall curve and coordinate axis is 0.9673, increased by 0.3%. The loss values of all models eventually converged between 0.4 and 0.5. The results show that compared with Mask-RCNN model, data augmentation can solve the model overfitting problem that is caused by the small amount of data. By fusing the multi-layer features of Citrus green fruits, the improved model can pay more attention to the obvious regions and more detailed features in each image, so as to improve the accuracy of citrus green fruit detection and segmentation.

Under the conditions of this study, the loss value of the ResNet50 network converged slightly faster than that of the ResNet101 network, and the model accuracy of the ResNet50 network was 1.93% higher than that of the ResNet101 network. The reason for this phenomenon is that the parameter scale of ResNet101 is larger than that of ResNet50, and the loss value is related to the update speed of the weight parameters. Compared with the multicategory and multi scene target detection task, the semantic information of the citrus green fruit images that were collected in this study is not obvious, and the low level of demand for rich semantic features that are generated at higher

TABLE 3 Areas that are enclosed by the PR curves and coordinate axes of the four models.

Models	Areas
Mask-RCNN+ResNet50	0.9638
Mask-RCNN+ResNet101	0.9549
Improved Mask-RCNN+ResNet50	**0.9673**
Improved Mask-RCNN+ResNet101	0.9556

Bold values represent the best indicators in the current table.

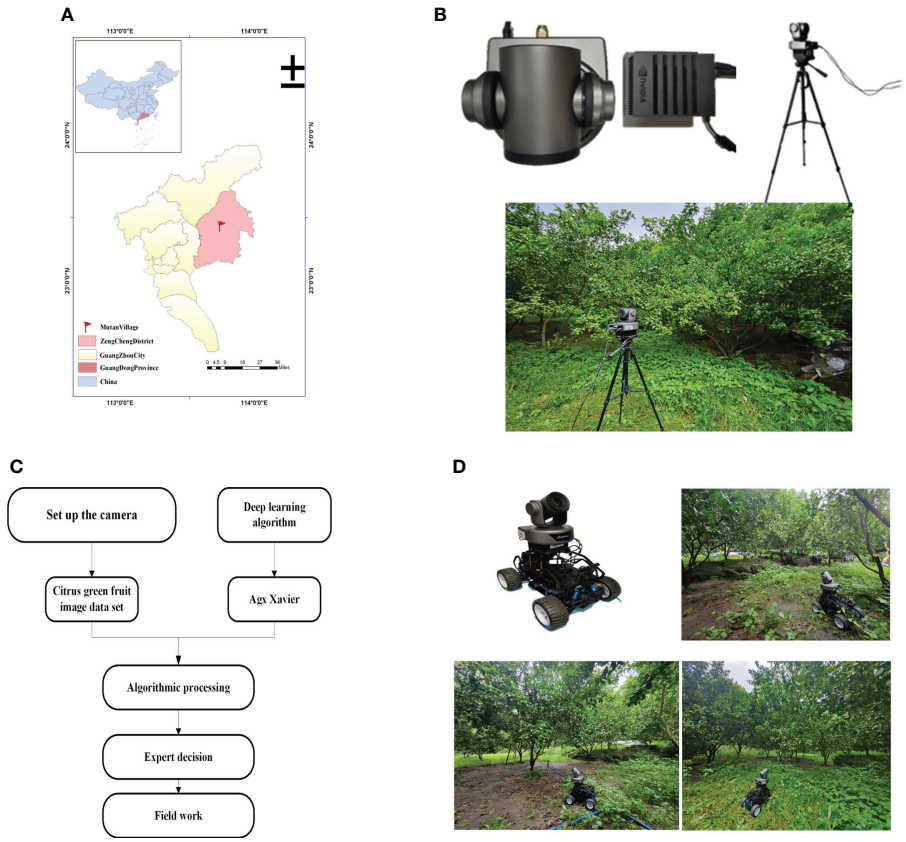


FIGURE 12
(A) Example application location; **(B)** HD camera, Agx Xavier, Overall build, Field layout effect;**(C)** Overall flow chart; **(D)** Unmanned vehicle and unmanned vehicle for citrus green fruit dynamic detection.

stages results in data-level limitations,the feature maps with too many semantic features results in model overfitting. Therefore, the improved Mask-RCNN +ResNet50 model had better inference results on the test set, thereby indicating that the model had more accurate fitting and better generalization ability for single citrus green fruit characteristics.

Based on the improved model proposed in this study, we carried out the assisted recognizing detection of citrus green fruit thinning. The results show that when the improved model proposed in this paper is applied in the field, the accuracy meets the training effect, and the average inference time of the system is 0.969s, which meets the requirements of real-time detection. The improved model proposed in this paper has good generalization ability and can be practically applied to the identification of citrus green fruits in citrus orchards. The inference results can further help the development of fruit thinning in orchards.

Based on the observations of this study, the designed model has various limitations. In a more complex citrus orchard scene, the overlap effect cannot be ignored when the number of recognized citrus green fruits is large and the pixel sizes do not differ much. In the next step of research, we need to further differentiate the analysis and training for different scenes, and the model will be lightly modified to further improve its generalization ability to enhance its detection accuracy and speed.

6 Conclusions

Using computer vision technologies to establish a citrus detection model and realize the real-time processing of orchard images through sensor systems and high-performance computers is an important development trend of smart agriculture in citrus detection at present. This study designed a citrus green fruit detection method based on improved feature network extraction. Citrus green fruits were selected as the research objects. To ensure the accurate detection of different individuals of the same target category in the collected images in the actual detection and recognition process, based on the Mask-RCNN algorithm, CB-Net with deep and shallow fusion was innovatively combined with the traditional feature extraction network ResNet to fuse the multilayer features of citrus green fruits. Then, the proposed model was compared with the traditional model in detail.

The improved model that was proposed in this study showed an average accuracy of 95.36 on the citrus green fruit dataset that was established in this study, and the accuracy of the model increased by 1.18% after the same feature extraction network was integrated into CB-Net. The improved model has higher ability to express the characteristics of citrus green fruits and reduces the interference of complex image backgrounds in citrus green fruit instance segmentation, and the detection accuracy has exceeded the basic accuracy requirements. Compared with the current common algorithms, the improved model has accuracy advantages and edge erection feasibility. In a data enhancement contrast test, it was verified that the data enhancement method that was proposed in this study can solve the overfitting problem that is caused by the limited amount of data. The results showed that the proposed algorithm based on the improved Mask-RCNN has improved citrus green fruit detection accuracy and speed. A performance comparison of the Mask-RCNN algorithm before and after the improvement showed that the optimal model for citrus green fruit detection and segmentation is the improved Mask-RCNN +ResNet50 model. The results showed that the ResNet50 network with low-level semantic information performed better on the dataset in this study. CB-Net promotes the extraction and fusion of features and further improves the generalization ability of the model to unknown data.

In the future, the citrus green fruit detection and segmentation model can be further applied to an edge intelligent platform with high computing power. The deployment of the model at the edge is another important strategy for citrus orchard data transmission. It can provide technical support for the management and control system of the citrus key growth period and provide a reference for the realization of scientific and unmanned farm construction.

Data availability statement

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

Author contributions

JLu conceptualized the experiments, selected the algorithms, collected and analyzed the data. RY, CY and JLi wrote the manuscript. HW and WC trained the algorithms and collected and analyzed the data. XC, WW and YL supervised the project and revised the manuscript. All authors discussed and revised the manuscript. All authors contributed to the article and approved the submitted version.

Funding

This work was supported by the Laboratory of Lingnan Modern Agriculture Project (NT2021009), Basic and Applied Basic Research Project of Guangzhou Basic Research Plan in 2022 (202201010077), The 111 Project (D18019), Guangzhou Key R&D project (SL2022B03J01345), The Open Research Fund of Guangdong Key Laboratory for New Technology Research of Vegetables (201704) and Guangdong Province Enterprise Science and Technology Special Ombudsman Project (GDKTP2020070200).

Conflict of interest

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

Publisher's note

References

Apolo-Apolo, O., Martínez-Guanter, J., Egea, G., Raja, P., and Pérez-Ruiz, M. (2020). Deep learning techniques for estimation of the yield and size of citrus fruits using a UAV. *Eur. J. Agron.* 115, 126030. doi: 10.1016/j.eja.2020.126030

Deng, Y., Wu, H., and Zhu, H. (2020). Recognition and counting of citrus flowers based on instance segmentation. *Trans. Chin. Soc Agric. Eng.* 36, 200–207. doi: 10.11975/j.issn.1002-6819.2020.07.023

Dorj, U.-O., Lee, M., and Yun, S.-S. (2017). An yield estimation in citrus orchards via fruit detection and counting using image processing. *Comput. Electron. Agric.* 140, 103–112. doi: 10.1016/j.compag.2017.05.019

Fan, P., Lang, G., Guo, P., Liu, Z., Yang, F., Yan, B., et al. (2021). Multi-feature patch-based segmentation technique in the Gray-centered RGB color space for improved apple target recognition. *Agriculture* 11, 273. doi: 10.3390/agriculture11030273

Fu, L., Feng, Y., Majeed, Y., Zhang, X., Zhang, J., Karkee, M., et al. (2018). Kiwifruit detection in field images using faster r-CNN with ZFNet. *IFAC-PapersOnLine* 51, 45–50. doi: 10.1016/j.ifacol.2018.08.059

Greener, J. G., Kandathil, S. M., Moffat, L., and Jones, D. T. (2022). A guide to machine learning for biologists. *Nat. Rev. Mol. Cell. Bio* 23, 40–55. doi: 10.1038/s41580-021-00407-0

Haseeb, K., Ud Din, I., Almogren, A., and Islam, N. (2020). An energy efficient and secure IoT-based WSN framework: An application to smart agriculture. *Sensors* 2081, 2081. doi: 10.3390/s20072081

He, K., Gkioxari, G., Dollár, P., and Girshick, R. (2017). "Mask r-cnn," in *Proceedings of the IEEE International Conference on Computer Vision (ICCV)*. 2961–2969 (New York: NY: IEEE). doi: 10.1109/ICCV.2017.322

He, Z., Xiong, J., Chen, S., Li, Z., Chen, S., Zhong, Z., et al. (2020). A method of green citrus detection based on a deep bounding box regression forest. *Biosyst. Eng.* 193, 206–215. doi: 10.1016/j.biosystemseng.2020.03.001

He, K., Zhang, X., Ren, S., and Sun, J. (2016). ""Deep residual learning for image recognition"," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*. 770–778 (Las Vegas: NV). doi: 10.1109/CVPR.2016.90

Hong, S.-J., Kim, S.-Y., Kim, E., Lee, C.-H., Lee, J.-S., Lee, D.-S., et al. (2020). Moth detection from pheromone trap images using deep learning object detectors. *Agriculture* 10, 170. doi: 10.3390/agriculture10050170

Hu, X. (2018). *Detecting green citrus fruits on the trees based on vision* (Huazhong Agricultural University).

Ioffe, S., and Szegedy, C. (2015). Batch normalization: Accelerating deep network training by reducing internal covariate shift. *arXiv e-prints arXiv* 1502, 03167. doi: 10.48550/arXiv.1502.03167

Jia, W., Zhang, Z., Shao, W., Hou, S., Ji, Z., Liu, G., et al. (2021).). FoveaMask: A fast and accurate deep learning model for green fruit instance segmentation. *Comput. Electron. Agric.* 191, 106488. doi: 10.1016/j.compag.2021.106488

Kurtulmus, F., Lee, W. S., and Vardar, A. (2011). Green citrus detection using 'eigenfruit', color and circular gabor texture features under natural outdoor conditions. *Comput. Electron. Agric.* 78, 140–149. doi: 10.1016/j.compag.2011.07.001

Lachat, E., Macher, H., Landes, T., and Grussenmeyer, P. (2015). Assessment and calibration of a RGB-d camera (Kinect v2 sensor) towards a potential use for close-range 3D modeling. *Remote Sens.* 7, 13070–13097. doi: 10.3390/rs71013070

Li, H., Chen, G., and Pei, A. (2020). Research on individual recognition of dairy cows based on improved mask r-CNN. *J. South. Chin. Agric. Univ.* 41, 161–168. doi: 10.7671/j.issn.1001-411X.202003030

Li, S., Hu, D., and Gao, S. (2019). Real-time classification and detection of citrus based on improved single short multibox detecter. *Trans. Chin. Soc Agric. Eng.* 35, 307–313. doi: 10.11975/j.issn.1002-6819.2019.24.036

Lin, T.-Y., Dollár, P., Girshick, R., He, K., Hariharan, B., and Belongie, S. (2017). "Feature pyramid networks for object detection," in *Proceedings of the IEEE conference on computer vision and pattern recognition*. 2117–2125 (Honolulu, HI: IEEE). doi: 10.1109/CVPR.2017.106

Linker, R., Cohen, O., and Naor, A. (2012). Determination of the number of green apples in RGB images recorded in orchards. *Comput. Electron. Agric.* 81, 45–57. doi: 10.1016/j.compag.2011.11.007

Lin, G., Tang, Y., Zou, X., Li, J., and Xiong, J. (2019). In-field citrus detection and localisation based on RGB-d image analysis. *Biosyst. Eng.* 186, 34–44. doi: 10.1016/j.biosystemseng.2019.06.019

Liu, Y., Wang, Y., Wang, S., Liang, T., Zhao, Q., Tang, Z., et al. (2020). "Cbnet: A novel composite backbone network architecture for object detection," in *Proceedings of the AAAI conference on artificial intelligence (New York)*, (Palo Alto, California USA: AAAI) Vol. 34. 11653–11660. doi: 10.1609/aaai.v34i07.6834

Liu, S., Yang, C., Hu, Y., Huang, L., and Xiong, L. (2018). ""A method for segmentation and recognition of mature citrus and branches-leaves based on regional features"," in *Chinese Conference on image and graphics technologies* (Singapore: Springer), 292–301.

Liu, B., Zhao, W., and Sun, Q. (2017). "Study of object detection based on faster r-CNN," in *2017 Chinese automation congress (CAC)* (Jinan, China: IEEE). doi: 10.1109/CAC.2017.8243900

Li, H., Zhang, M., Gao, Y., Li, M., and Ji, Y. (2017). Green ripe tomato detection method based on machine vision in greenhouse. *Trans. Chin. Soc Agric. Eng.* 33, 328–334. doi: 10.11975/j.issn.1002-6819.2017.z1.049

Lu, J., and Hu, X. (2017). Detecting green citrus fruit on trees in low light and complex background based on MSER and HCA. *Trans. Chin. Soc Agric. Eng.* 33, 196–201. doi: 10.11975/j.issn.1002-6819.2017.19.025

Peng, Y., Zhao, S., and Liu, J. (2021). Fused deep features-based grape varieties identification using support vector machine. *Agriculture* 11, 869. doi: 10.3390/agriculture11090869

Qassim, H., Verma, A., and Feinzimer, D. (2018). "Compressed residual-VGG16 CNN model for big data places image recognition," in *2018 IEEE 8th annual computing and communication workshop and conference (CCWC)* (Las Vegas: NV: IEEE), 169–175. doi: 10.1109/CCWC.2018.8301729

Rahnemoonfar, M., and Sheppard, C. (2017). Deep count: fruit counting based on deep simulated learning. *Sensors* 17, 905. doi: 10.3390/s17040905

Rakun, J., Stajnko, D., and Zazula, D. (2011). Detecting fruits in natural scenes by using spatial-frequency based texture analysis and multiview geometry. *Comput. Electron. Agric.* 76, 80–88. doi: 10.1016/j.compag.2011.01.007

Ren, S., He, K., Girshick, R., and Sun, J. (2015). Faster r-cnn: Towards real-time object detection with region proposal networks. *Adv. Neural Inf. Process. Syst.* 2015, 28. doi: 10.1109/TPAMI.2016.2577031

Russell, B. C., Torralba, A., Murphy, K. P., and Freeman, W. T. (2008). Labelme: a database and web-based tool for image annotation. *Int. J. Comput. Vision.* 77, 157–173. doi: 10.1007/s11263-007-0090-8

Sa, I., Ge, Z., Dayoub, F., Upcroft, B., Perez, T., and McCool, C. (2016). Deepfruits: A fruit detection system using deep neural networks. *Sensors* 1222, 1222. doi: 10.3390/s16081222

Santos, T. T., de Souza, L. L., dos Santos, A. A., and Avila, S. (2020). Grape detection, segmentation, and tracking using deep neural networks and three-dimensional association. *Comput. Electron. Agric.* 170, 105247. doi: 10.1016/j.compag.2020.105247

Shilei, L., Sihua, L., Zhen, L., Tiansheng, H., Yueju, X., and Benlei, W. (2019). Orange recognition method using improved YOLOv3-LITE lightweight neural network. *Trans. Chin. Soc Agric. Eng.* 35, 10. doi: 10.11975/j.issn.1002-6819.2019.17.025

Wahabzada, M., Paulus, S., Kersting, K., and Mahlein, A.-K. (2015). Automated interpretation of 3D laserscanned point clouds for plant organ segmentation. *BMC Bioinf.* 16, 1–11. doi: 10.1186/s12859-015-0665-2

Wang, L., Zhao, Y., Liu, S., Li, Y., Chen, S., and Lan, Y. (2022). Precision detection of dense plums in orchards using the improved YOLOv4 model. *Frontiers in Plant Science* 13, 839269.

Wang, D., and He, D. (2019). Recognition of apple targets before fruits thinning by robot based on r-FCN deep convolution neural network. *Trans. Chin. Soc Agric. Eng.* 35, 156–163. doi: 10.11975/j.issn.1002-6819.2019.03.020

Wang, Y., Lv, J., Xu, L., Gu, Y., Zou, L., and Ma, Z. (2020). A segmentation method for waxberry image under orchard environment. *Sci. Hortic.* 266, 109309. doi: 10.1016/j.scienta.2020.109309

Wang, C., and Xiao, Z. (2021). Potato surface defect detection based on deep transfer learning. *Agriculture* 11, 863. doi: 10.3390/agriculture11090863

Wang, C., Zou, X., Tang, Y., Luo, L., and Feng, W. (2016). Localisation of litchi in an unstructured environment using binocular stereo vision. *Biosyst. Eng.* 145, 39–51. doi: 10.1016/j.biosystemseng.2016.02.004

Yamamoto, K., Guo, W., Yoshioka, Y., and Ninomiya, S. (2014). On plant detection of intact tomato fruits using image analysis and machine learning methods. *Sensors* 14, 12191–12206. doi: 10.3390/s140712191

Yang, X., Shu, L., Chen, J., Ferrag, M. A., Wu, J., Nurellari, E., et al. (2021). A survey on smart agriculture: Development modes, technologies, and security and privacy challenges. *IEEE/CAA J. Autom. Sinica.* 8, 273–302. doi: 10.1109/JAS.2020.1003536

Yan, H., Liu, C., Zhao, J., Ye, X., Wu, Q., Yao, T., et al. (2021). Genome-wide analysis of the NF-y gene family and their roles in relation to fruit development in tartary buckwheat (Fagopyrum tataricum). *Int. J. Biol. MACROMOL.* 190, 487–498. doi: 10.1016/j.ijbiomac.2021.09.001

Yin, H., Gu, Y. H., Park, C.-J., Park, J.-H., and Yoo, S. J. (2020). Transfer learning-based search model for hot pepper diseases and pests. *Agriculture* 10, 439. doi: 10.3390/agriculture10100439

Yoosefzadeh Najafabadi, M. (2021). *Using advanced proximal sensing and genotyping tools combined with bigdata analysis methods to improve soybean yield* (University of Guelph). Ph.D. thesis.

Zhang, Y. (2020). *Research on apple picking of robot arm based on deep learning* (Beijing Forestry University). Master's thesis.

Zhao, C., Lee, W. S., and He, D. (2016). Immature green citrus detection based on colour feature and sum of absolute transformed difference (SATD) using colour images in the citrus grove. *Comput. Electron. Agric.* 124, 243–253. doi: 10.1016/j.compag.2016.04.009

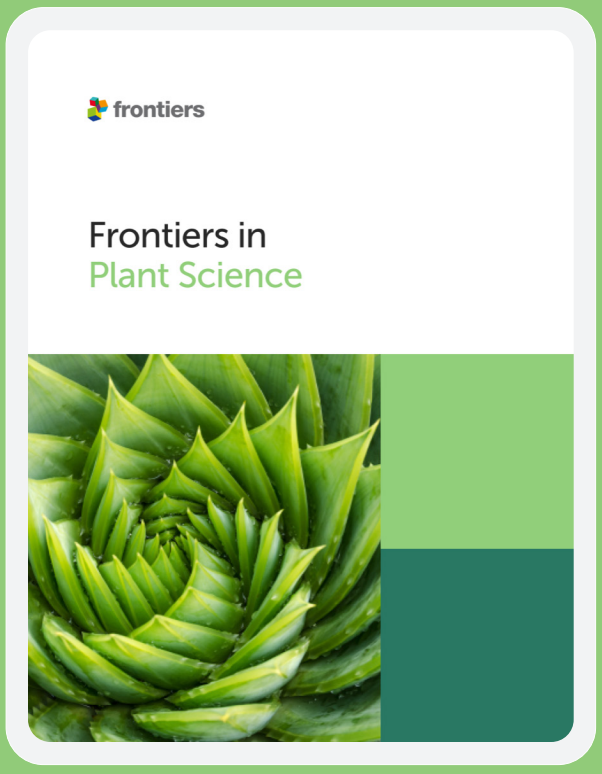
Frontiers in
Plant Science

Cultivates the science of plant biology and its applications

The most cited plant science journal, which advances our understanding of plant biology for sustainable food security, functional ecosystems and human health.

Discover the latest Research Topics

See more →



Frontiers

Avenue du Tribunal-Fédéral 34
1005 Lausanne, Switzerland
frontiersin.org

Contact us

+41 (0)21 510 17 00
frontiersin.org/about/contact