INTELLIGENT COMPUTING RESEARCH WITH APPLICATIONS IN ECOLOGICAL PLANT PROTECTION

EDITED BY: Yu Xue, Jian Su, Ferrante Neri and Ali Wagdy Mohamed

PUBLISHED IN: Frontiers in Plant Science

frontiers Research Topics

frontiers

ISSN 1664-8714
ISBN 978-2-83250-899-2
DOI 10.3389/978-2-83250-899-2

About Frontiers

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

Frontiers Journal Series

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

Dedication to Quality

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

What are Frontiers Research Topics?

Frontiers Research Topics are very popular trademarks of the Frontiers Journals Series: they are collections of at least ten articles, all centered on a particular subject. With their unique mix of varied contributions from Original Research to Review Articles, Frontiers Research Topics unify the most influential researchers, the latest key findings and historical advances in a hot research area! Find out more on how to host your own Frontiers Research Topic or contribute to one as an author by contacting the Frontiers Editorial Office: frontiersin.org/about/contact

INTELLIGENT COMPUTING RESEARCH WITH APPLICATIONS IN ECOLOGICAL PLANT PROTECTION

Topic Editors:
Yu Xue, Nanjing University of Information Science and Technology, China
Jian Su, Nanjing University of Information Science and Technology, China
Ferrante Neri, University of Surrey, United Kingdom
Ali Wagdy Mohamed, Cairo University, Egypt

Citation: Xue, Y., Su, J., Neri, F., Mohamed, A. W., eds. (2022).
Intelligent Computing Research with Applications in Ecological Plant Protection.
Lausanne: Frontiers Media SA. doi: 10.3389/978-2-83250-899-2

Table of Contents

frontiers in Plant Science

ORIGINAL RESEARCH
published: 28 February 2022
doi: 10.3389/fpls.2022.849606

Green Visual Sensor of Plant: An Energy-Efficient Compressive Video Sensing in the Internet of Things

Ran Li[1]*, *Yihao Yang*[1] *and Fengyuan Sun*[2]

[1] School of Computer and Information Technology, Xinyang Normal University, Xinyang, China, [2] Guangxi Key Laboratory of Wireless Wideband Communication and Signal Processing, Guilin University of Electronic Technology, Guilin, China

Internet of Things (IoT) realizes the real-time video monitoring of plant propagation or growth in the wild. However, the monitoring time is seriously limited by the battery capacity of the visual sensor, which poses a challenge to the long-working plant monitoring. Video coding is the most consuming component in a visual sensor, it is important to design an energy-efficient video codec in order to extend the time of monitoring plants. This article presents an energy-efficient Compressive Video Sensing (CVS) system to make the visual sensor green. We fuse a context-based allocation into CVS to improve the reconstruction quality with fewer computations. Especially, considering the practicality of CVS, we extract the contexts of video frames from compressive measurements but not from original pixels. Adapting to these contexts, more measurements are allocated to capture the complex structures but fewer to the simple structures. This adaptive allocation enables the low-complexity recovery algorithm to produce high-quality reconstructed video sequences. Experimental results show that by deploying the proposed context-based CVS system on the visual sensor, the rate-distortion performance is significantly improved when comparing it with some state-of-the-art methods, and the computational complexity is also reduced, resulting in a low energy consumption.

Keywords: Internet of Things, visual sensor, Compressive Video Sensing, context extraction, linear recovery, plant monitoring

OPEN ACCESS

Edited by:
Yu Xue,
Nanjing University of Information Science and Technology, China

Reviewed by:
Romany Mansour,
The New Valley University, Egypt
Zijian Qiao,
Ningbo University, China
Khan Muhammad,
Sejong University, South Korea

***Correspondence:**
Ran Li
liran@xynu.edu.cn

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

Received: *06 January 2022*
Accepted: *24 January 2022*
Published: *28 February 2022*

Citation:
Li R, Yang Y and Sun F (2022) Green Visual Sensor of Plant: An Energy-Efficient Compressive Video Sensing in the Internet of Things. Front. Plant Sci. 13:849606. doi: 10.3389/fpls.2022.849606

1. INTRODUCTION

In the Internet of Things (IoT), the plant propagation process or plant growth can be monitored by visual sensors. One benefit from the framework of IoT, a large amount of data on the plant can be gathered in a central server, and the valuable information can be achieved by analyzing the data in real-time. However, with the limited processing capabilities and power/energy budget of visual sensors, it is a challenge for video monitoring of plant to compress large-scale video sequences by using the traditional codec, e.g., H.264/AVC and HEVC (Sullivan et al., 2012), so the existing works have developed low-complexity and energy-efficient video codecs, in which Distributed Video Coding (DVC) (Girod et al., 2005) and Compressive Video Sensing (CVS) (Baraniuk et al., 2017) have attracted more attention in industry and academia. Different from DVC, CVS dispenses with the feedback and virtual channels (Unde and Pattathil, 2020), which makes the codec framework simpler. Meanwhile, CVS provides a low-complexity encoder because of its theoretic foundation, Compressive Sensing (CS) (Baraniuk, 2007), realizes the capture of

video frames at a rate significantly below the Nyquist rate. Currently, many researchers recognize that CVS is a potential scheme to compress the video sequences in the IoT framework, and especially for wireless video monitoring of plants, the CVS scheme can assist visual sensors to efficiently reduce the energy consumptions, however, its rate-distortion performances are still far from satisfactory.

The objective of this article is to improve the rate-distortion performance of CVS, providing high-quality video monitoring of plants with low energy consumption. To achieve this objective, the existing works focus on how to design excellent recovery algorithms, and they are keen on mixing various advanced tools into the CVS framework, e.g., the latest popular Deep Neural Network (DNN) (Palangi et al., 2016; Zhao et al., 2020; Tran et al., 2021). Though effective, they bear a heavy computational burden. Different from these works, we try to exploit the capability of CS to capture important structures, improving the reconstruction quality only armed with some simple recovery algorithms. It is well known that the context feature (Shechtman and Irani, 2007; Romano and Elad, 2016) is a good structure for visual quality, and, therefore, in this article, we focus on how to fuse contexts into CVS for an obvious improvement of reconstruction quality.

Compressive Video Sensing consists of three essential steps including CS measurement, measurements quantization, and reconstruction. CS measurement is a process of randomly sampling each video frame, in which the block-based (Gan, 2007; Bigot et al., 2016) or structurally (Do et al., 2012; Zhang et al., 2015) random matrix is often used to ensure the small memory requirement. Output by CS measurement, all measurements are required to be quantized as bits, then transmitted to the decoder. The straightforward solution to incorporating quantization into CVS is simply to apply Scalar Quantization (SQ), but it brings a big error. For block-based sampling, Differential Pulse Code Modulation (DPCM) (Mun and Fowler, 2012) can be used, and it exploits the correlations between blocks to improve the rate-distortion performance. Based on DPCM, many works also proposed some efficient predictive schemes (Zhang et al., 2013; Gao et al., 2015) to quantize CS measurements. Reconstruction is deployed at the decoder, and it uses quantized measurements to reconstruct the video sequence by the CS recovery algorithm. At present, the reconstruction can be implemented by one of the three types: frame-by-frame (Chen Y. et al., 2020; Trevisi et al., 2020), three-dimensional (3D) (Qiu et al., 2015; Tachella et al., 2020), and distributed strategies (Zhang et al., 2020; Zhen et al., 2020). The frame-by-frame reconstruction performs a CS recovery algorithm to reconstruct each video frame independently, and it has a poor rate-distortion performance due to neglecting the correlations between frames. The 3D reconstruction designs some complex representation models to once reconstruct a whole video sequence or a Group Of Pictures (GOP), e.g., Li et al. (2020) proposed the Scalable Structured CVS (SS-CVS) framework, which learns the union of data-driven subspaces model to reconstruct GOPs. However, it has a defect in 3D reconstruction that the huge memory and high computational complexity are required to be invested at decoder. Derived from the decoding strategy of DVC, the distributed reconstruction divides the input video sequence into non-key frames and key frames and reconstructs each non-key frame by the CS recovery algorithm with the aid of its neighboring key frames. With a small memory and a low computational complexity, the distributed reconstruction improves the rate-distortion performance by exploiting the motions between frames, so many existing works focus on it to design the CVS systems, e.g., Ma et al. proposed the DIStributed video Coding Using Compressed Sampling (DISCUCS) (Prades-Nebot et al., 2009), Gan et al. proposed the DIStributeCOmpressed video Sensing (DISCOS) (Do et al., 2009), Fowler et al. proposed the Multi-Hypothesis Block CS (MH-BCS) system (Chen et al., 2011; Tramel and Fowler, 2011; Azghani et al., 2016), etc. The core of distributed reconstruction is the Multi-Hypothesis (MH) predictive technique, which uses a linear combination of blocks in key frames to interpolate the blocks in non-key frames. As one of the state-of-the-art techniques, the MH prediction is widely applied to distributed reconstruction. Recently, some works try to modify the implementation of MH prediction, e.g., Chen C. et al. (2020) added the iterative Reweighted TIKhonov-regularized scheme into MH prediction (MH-RTIK), causing a significant improvement of CVS performance. CS theory indicates that the precise recovery requires enough CS measurements. With insufficient CS measurements, the excellent CS recovery algorithm still cannot prevent the degradation of reconstruction quality, however, by adaptively allocating CS measurements based on local structures of the image, a simple recovery algorithm can also provide a good reconstruction quality (Yu et al., 2010; Taimori and Marvasti, 2018; Zammit and Wassell, 2020). Judging from the above facts, the adaptive allocation is a potential way to improve the rate-distortion performance of the CVS system with a light codec.

This article presents a context-based CVS system, of which the core is the allocation of CS measurements adapted by context structures at the encoder. Based on these adaptive measurements, by combining linear estimation and MH prediction into distributed reconstruction, the decoder provides a satisfying reconstruction quality with low memory and computational cost. The contributions of the proposed context-based CVS is to solve the following issues:

(1) How to extract the context structures from CS measurements? Traditional methods use pixels to compute the context features, but it costs lots of computations at the encoder, resulting in impracticality for CVS. Especially when the encoder is realized by Compressive Imaging (CI) devices (Liu et al., 2019; Deng et al., 2021), due to the unavailability of original pixels, it is impossible to perform the traditional methods. Considering the low dimensionality and availability of CS measurements, it is practical in CVS to extract context structures from CS measurements.
(2) How to adaptively allocate CS measurements by context structures? Contexts measure the correlations between pixels, and their distribution reveals some meaningful structures, e.g., smoothness, edges, textures, etc. With the same recovery quality, fewer necessary measurements are required for simple structures and more for complex structures. According to the distribution of contexts, an efficient

allocation is designed to avoid insufficiency or redundancy of measurements.

(3) How to quantize the adaptive measurements? Adaptive allocation makes blocks have different numbers of CS measurements, as a result, the traditional prediction cannot be applied to quantization. Due to the insufficient capability of SQ, an appropriate prediction scheme is required to reduce the quantization error.

Experimental results show that the proposed context-based CVS system outputs the high-quality reconstructed video sequences when monitoring plant growth or propagation and improves the rate-distortion performances when compared with the state-of-the-art CVS systems, which demonstrates the effectiveness of context-based allocation for CVS.

The rest of this article is organized as follows. Section 2 briefly overviews Plant Monitoring System, CVS, and describes the traditional method to extract context features. Section 3 presents the proposed context based CVS system. Experimental results are provided in Section 4, and we conclude this article in Section 5.

2. RELATED WORKS

2.1. Plant Monitoring System

In modern agriculture, it is essential to monitor plant propagation or growth for guaranteeing productivity. The labor costs can be efficiently reduced by automatically capturing the architectural parameters of the plant, so more and more attention has been paid to the design of the plant monitoring system (Somov et al., 2018; Grimblatt et al., 2021; Rayhana et al., 2021). Early, lots of systems are designed to monitor the various environmental parameters on plant growth, such as humidity, temperature, solar illuminance, etc., e.g., James and Maheshwar (2016) used multiple sensors to measure the soil data of plants and transmitted these data to the mobile phone by Raspberry Pi; Okayasu et al. (2017) developed a self-powered wireless monitoring device that is equipped with some environmental sensors; Guo et al. (2018) added big-data services to analyze the environmental data on plant growth. These environmental parameters indirectly indicate the process of plant growth, and they cannot record the visual scenes on plant growth, resulting in the unavailability of the physical structure parameters on plants. To realize the visual monitoring of plants, some works have started to integrate the visual sensors into the plant monitoring system, e.g., Peng et al. (2022) used the binocular camera to capture video sequences on a plant and used the structure from motion method (Piermattei et al., 2019) to extract the 3-D information of a plant; Sajith et al. (2019) designed a complex network to derive the plant growth parameters from the monitoring images; Akila et al. (2017) extracted the plant color and texture by the visual monitoring system. From the above, it can be seen that the visual sensor or camera is used to capture the video sequences on plant growth, and these video sequences are compressed as bitstream which is transmitted to the IoT cloud for further analyzing. As the core of visual sensors, the video compression is a major energy consumer, so a challenge that we face for the visual monitoring system of the plant is to design an energy-efficient video coding scheme to extend the working time of the visual sensor. In the framework of IoT, CVS is a potential coding scheme to reduce the energy consumption of visual sensors. The following briefly overviews the CVS systems.

2.2. CVS System

Compressive Video Sensing is the marriage of CS theory and DVC, which reduces the encoding costs and enhances the robustness to noise, thus becoming a potential video codec for wireless visual sensors. At the encoder, to satisfy low complexity and fast computation, the block-based CS sampling is performed on each video frame independently, i.e., the ith video frame $\boldsymbol{f}_i$ of size $N_1 \times N_2$ is partitioned into non-overlapping blocks of size $B \times B$, each block is vectorized as $\boldsymbol{x}_{i,j}$ of length N_b, and the CS measurements $\boldsymbol{y}_{i,j}$ of $\boldsymbol{x}_{i,j}$ are output by

$$\boldsymbol{y}_{i,j} = \boldsymbol{\Phi}_{i,j} \cdot \boldsymbol{x}_{i,j} \tag{1}$$

where $\boldsymbol{\Phi}_{i,j}$ is called as the measurement matrix and can be constructed by some random matrices, e.g., Gaussian, Bernoulli, structural random matrix, etc. By setting the length of $\boldsymbol{y}_{i,j}$ to be $M_{i,j}$, the size of $\boldsymbol{\Phi}_{i,j}$ is fixed to be $M_{i,j} \times N_b$, and the subrate S_i of $\boldsymbol{f}_i$ is defined as

$$S_i = \frac{M_i}{N} = \frac{\sum_{j=1}^{J} M_{i,j}}{N_1 \times N_2} \tag{2}$$

where N is the number of total pixels in $\boldsymbol{f}_i$, M_i is the number of CS measurements for $\boldsymbol{f}_i$, and J is the number of blocks in $\boldsymbol{f}_i$. In CI application, an optical device is designed to perform Equation (1), and directly output the CS measurements. To ensure a stable recovery, L video frames are gathered to form a GOP, in which the first frame, called the key frame, is set to be a high subrate, and others, called the non-key frame, are set to be a low subrate. After quantization, all CS measurements of GOP are packaged and transmitted to decoder.

At the decoder, by using the received CS measurements, the frame-by-frame, 3D, or distributed strategy is performed to reconstruct the GOP. For frame-by-frame, the reconstruction model can be represented by

$$\left\{\hat{\boldsymbol{x}}_{i,j}\right\}_{j=1}^{J} = arg \min_{\left\{\boldsymbol{x}_{i,j}\right\}_{j=1}^{J}} \left\{ \sum_{j=1}^{J} \left\| \boldsymbol{y}_{i,j} - \boldsymbol{\Phi}_{i,j} \cdot \boldsymbol{x}_{i,j} \right\|_2^2 + \alpha \cdot \sum_{j=1}^{J} \left\| \boldsymbol{\Psi} \cdot \boldsymbol{x}_{i,j} \right\|_1 \right\} \tag{3}$$

where $\boldsymbol{\Psi}$ denotes the 2D sparse representation basis, α is a regularization factor, $\|\cdot\|_2$ denotes ℓ_2 norm, and $\|\cdot\|_1$ denotes ℓ_1 norm. The model (3) can be solved by some non-linear optimization algorithms, e.g., Alternating Direction Method of Multipliers (ADMM) (Yang et al., 2020), and all reconstructed blocks are spliced into the estimated frame $\hat{\boldsymbol{f}}_i$. The frame-by-frame model uses only the spatial correlations, so its rate-distortion performance is unsatisfactory. The 3D reconstruction

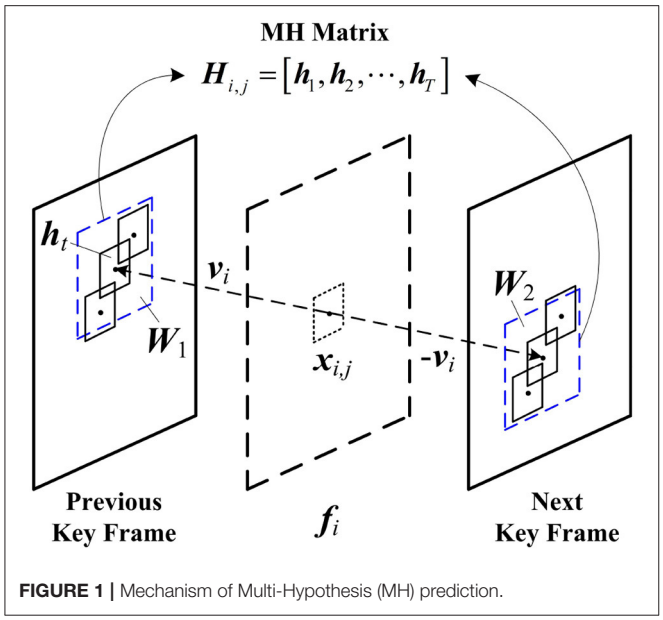


FIGURE 1 | Mechanism of Multi-Hypothesis (MH) prediction.

model fully considers the spatial-temporal correlations and it can be represented by

$$\left\{\hat{\boldsymbol{x}}_{i,j}\,|_{i=1}^{L}\right\}_{j=1}^{J} = arg \min_{\left\{\boldsymbol{x}_{i,j}|_{i=1}^{L}\right\}_{j=1}^{J}} \left\{ \sum_{i=1}^{L}\sum_{j=1}^{J} \|\boldsymbol{y}_{i,j} - \boldsymbol{\Phi}_{i,j}\cdot\boldsymbol{x}_{i,j}\|_2^2 + \alpha \cdot \sum_{i=1}^{L}\sum_{j=1}^{J} \left|\boldsymbol{\Gamma}\cdot\left[\boldsymbol{x}_{1,j}, \boldsymbol{x}_{2,j}, \cdots, \boldsymbol{x}_{L,j}\right]\right| \right\} \tag{4}$$

where $\boldsymbol{\Gamma}$ denotes the 3D sparse representation basis, and it is used to remove the spatial-temporal redundancies between blocks. Though effective, model (4) results in a heavy computational burden. Different from the 3D reconstruction, the distributed reconstruction uses the motion-compensation based prediction technique to expose the spatial-temporal redundancies between blocks. **Figure 1** shows the mechanism of MH prediction, which is commonly used in distributed reconstruction. MH prediction collects the spatial-temporal neighboring blocks in key frames to construct an MH matrix $\boldsymbol{H}_{i,j}$. According to the motion vector $\boldsymbol{v}_i$ of $\boldsymbol{x}_{i,j}$, the motion-aligned windows $\boldsymbol{W}_1$ and $\boldsymbol{W}_2$ of sizes $W \times W$ are, respectively, located on the previous and the next key frames, and all candidate blocks in $\boldsymbol{W}_1$ and $\boldsymbol{W}_2$ are extracted as the hypotheses $\{\boldsymbol{h}_t\}_{t=1}^{T}$ of $\boldsymbol{x}_{i,j}$, producing $\boldsymbol{H}_{i,j} = [\boldsymbol{h}_1, \boldsymbol{h}_2, \cdots, \boldsymbol{h}_T]$, in which $T = W^2$. By using MH prediction, the distributed reconstruction is modeled as a Least-Squares (LS) problem as follows:

$$\hat{\boldsymbol{w}}_{i,j} = arg\min_{\boldsymbol{w}} \left\{ \|\boldsymbol{y}_{i,j} - \boldsymbol{\Phi}_{i,j}\cdot\boldsymbol{H}_{i,j}\cdot\boldsymbol{w}\|_2^2 + \beta\cdot\|\boldsymbol{\Theta}\cdot\boldsymbol{w}\|_2^2 \right\} \tag{5}$$

$$\hat{\boldsymbol{x}}_{i,j} = \boldsymbol{H}_{i,j}\cdot\hat{\boldsymbol{w}}_{i,j} \tag{6}$$

where $\boldsymbol{\Theta}$ is the Tikhonov matrix, and β is a regularization factor. $\boldsymbol{\Theta}$ is a diagonal matrix and constructed by

$$\boldsymbol{\Theta} = \begin{bmatrix} \|\boldsymbol{y}_{i,j} - \boldsymbol{\Phi}_{i,j}\cdot\boldsymbol{h}_1\|_2 & & 0 \\ & \ddots & \\ 0 & & \|\boldsymbol{y}_{i,j} - \boldsymbol{\Phi}_{i,j}\cdot\boldsymbol{h}_T\|_2 \end{bmatrix} \tag{7}$$

With this structure, $\boldsymbol{\Theta}$ assigns weights of small magnitude to hypotheses mostly dissimilar from $\boldsymbol{x}_{i,j}$. The LS problem can be fast solved by the Conjugate Gradient algorithm (Zhang et al., 2018), which significantly reduces the computational complexity of distributed reconstruction. Due to the full exploitation of spatial-temporal correlations between blocks, the MH prediction enables the distributed reconstruction to provide superior recovery. From the above, in order to realize a light decoder and ensure a good recovery at the same time, distributed reconstruction is a wise way.

2.3. Contexts

Compressive Sensing theory indicates that the sparsity K of the signal determines its required number M of CS measurements by precise recovery. An empirical rule (Becker and Bobin, 2011) is that the precise recovery can be achieved if

$$M \geq 4\cdot K \tag{8}$$

In the block-based CS sampling, this rule can be used to avoid the redundancy or insufficiency of CS measurements for blocks, i.e., adapted by the sparsity, each block is allocated to the appropriate number of CS measurements. The sparsity is defined as the number of coefficients with significant magnitude in a representation, and its calculation has not a strict mathematical formula. For images, the sparsity can be revealed by some features, e.g., edge, variance, gradient, etc., and these features are applied into adaptive allocation, leading to the improvement of recovery quality. The simple features only describe the correlations between pixels, but the structures of blocks are not taken into consideration, thus we require some complex features to improve the efficiency of adaptive allocation. In Ref. Romano and Elad (2016), the self-similarity descriptor (Shechtman and Irani, 2007) is used to extract the contexts of blocks, which represents how similar a central block is to its large surrounding windows. Contexts contain the internal structures and external relations among blocks, and it is a potential feature to better reveal the sparsity variation. The following briefly describes how to extract the contexts in an image.

The context feature expresses the similarities between a central block and those of its large surrounding windows. As illustrated in **Figure 2**, for a central block $\boldsymbol{x}_p$ in an image, its similarity weights are computed by

$$s_{p,q} = \exp\left\{-\frac{\|\boldsymbol{x}_p - \boldsymbol{x}_q\|_2^2}{2\sigma^2}\right\}, \forall q \in \Omega_d\left(p\right) \tag{9}$$

where $\boldsymbol{x}_q$ denotes the qth surrounding block in a neighborhood $\Omega_d\left(p\right)$ of size $d \times d$, and σ is a normalization factor. The range

of $s_{p,q}$ is [0, 1], in which a large value indicates that the blocks $\boldsymbol{x}_p$ and $\boldsymbol{x}_q$ are highly similar, and a small value indicates that the two are substantially different. All weights constitute a correlation surface $\boldsymbol{U}_p = [s_{p,q}|\forall q \in \Omega_d(p)]$, of which the statistics reveal the self-similarity of $\boldsymbol{x}_p$. To measure the statistics, the correlation surface of $\boldsymbol{x}_p$ is rearranged into a histogram of b bins, of which the normalization is regarded as the context feature $\boldsymbol{g}_p$ of $\boldsymbol{x}_p$.

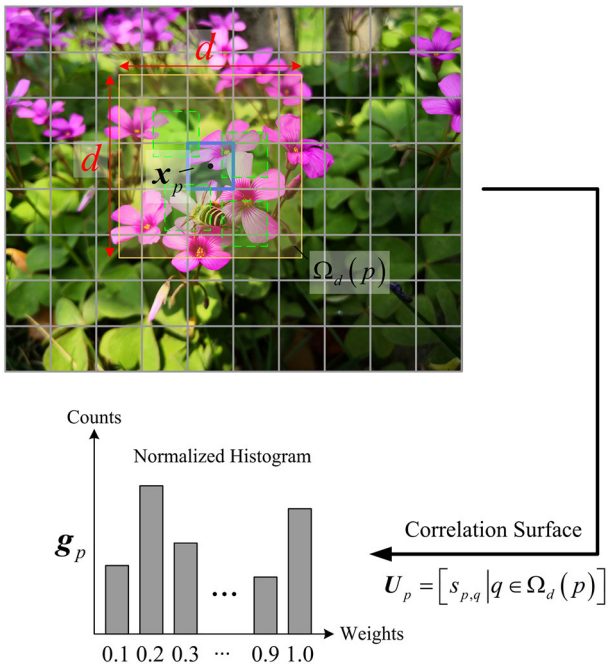


FIGURE 2 | Illustration on the traditional extraction of contexts.

The context feature $\boldsymbol{g}_p$ is an empirical distribution of the co-occurrences of $\boldsymbol{x}_p$ in its large surroundings, which measures the correlations between $\boldsymbol{x}_p$ to its surroundings. When $\boldsymbol{g}_p$ is biased toward the left bins, it can be concluded that the majority of $s_{p,q}$ are small, indicating the block $\boldsymbol{x}_p$ is unique, i.e., it originates from a highly textured and non-repetitive area, so its sparsity is relatively high. When $\boldsymbol{g}_p$ is biased toward the right bins, it means that most of $s_{p,q}$ are high, indicating that the block $\boldsymbol{x}_p$ has many co-occurrences in its surroundings, i.e., it originates from a large flat area, so its sparsity is low. From the above, we can see that the context feature accurately describes the geometric structure of a block with respect to its surrounding blocks, thus it is naturally sensitive to the sparsity variation. However, in CVS, the traditional method is impractical due to the unavailability of original pixels or high computational complexity. Therefore, it is challenging to extract the context feature by using CS measurements of blocks.

3. PROPOSED CONTEXT BASED CVS SYSTEM

3.1. System Architecture

As shown in **Figure 3**, we describe the architecture of the proposed context-based CVS system in detail. The input video sequence is divided into several GOPs of length L, and each GOP_k is successively encoded as Packet_k. After receiving this packet, the decoder reconstructs the corresponding $\widehat{\text{GOP}}_k$, and all reconstructed GOPs are regrouped as the entire video sequence.

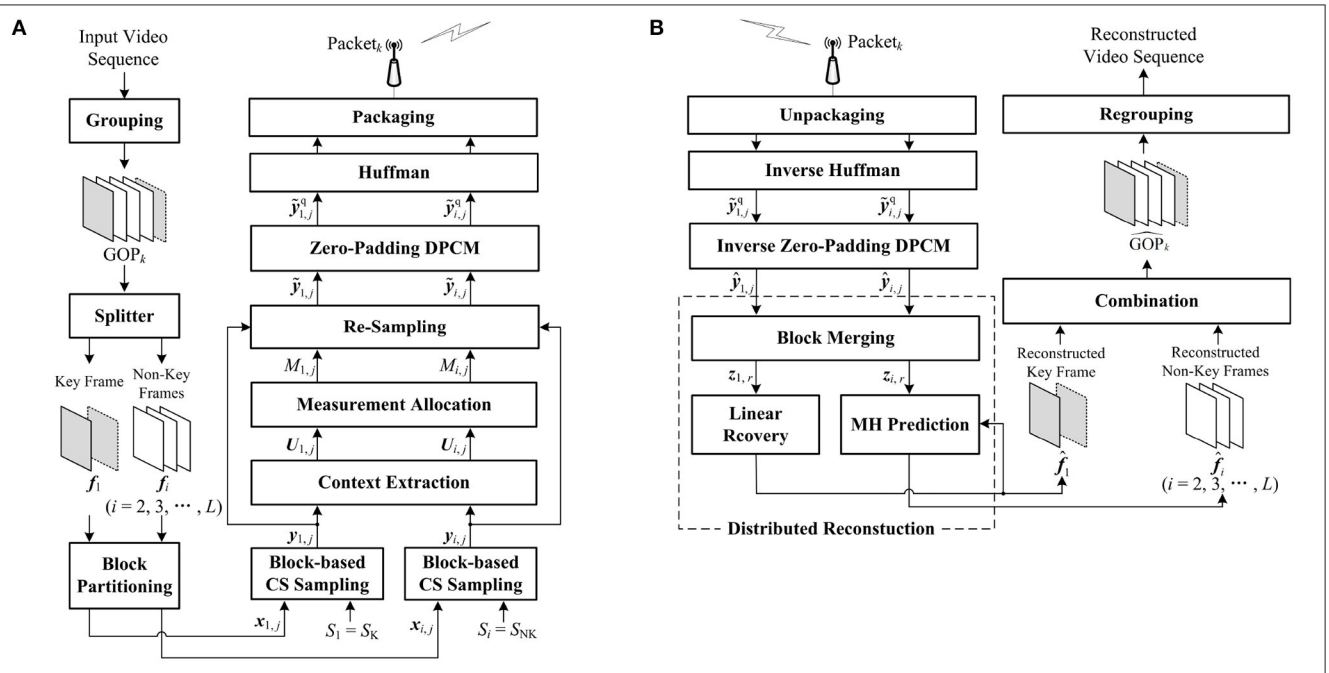


FIGURE 3 | Architecture of the proposed context-based Compressive Video Sensing (CVS) system: **(A)** encoder framework, **(B)** decoder framework.

Figure 3A presents the process of encoding GOP$_k$. The key frame $\boldsymbol{f}_1$ is split from GOP$_k$, and others $\{\boldsymbol{f}_i\}_{i=2}^{L}$ are regarded as the non-key frames. The key frame $\boldsymbol{f}_1$ and the ith non-key frame $\boldsymbol{f}_i$ are partitioned into J non-overlapping blocks $\{\boldsymbol{x}_{1,j}\}_{j=1}^{J}$ and $\{\boldsymbol{x}_{i,j}\}_{j=1}^{J}$ of size $B \times B$, respectively. For the key frame $\boldsymbol{f}_1$, we set a high subrate $S_1 = S_K$ to sample the blocks $\{\boldsymbol{x}_{1,j}\}_{j=1}^{J}$ and generate the CS measurements $\{\boldsymbol{y}_{1,j}\}_{j=1}^{J}$ according to Equation (1). The blocks $\{\boldsymbol{x}_{i,j}\}_{j=1}^{J}$ in the non-key frame $\boldsymbol{f}_i$ are sampled at a low subrate $S_i = S_{NK}$, producing the corresponding CS measurements $\{\boldsymbol{y}_{i,j}\}_{j=1}^{J}$ by Equation (1). For $\boldsymbol{f}_1$ and $\boldsymbol{f}_i$, based on the preset subrates, CS measurements are uniformly allocated to each block, however, without considering the structures of blocks, the uniform allocation results in either redundancy or insufficiency of CS measurements for some blocks. To improve the efficiency of block-based CS sampling, the core of the encoder is to perform the adaptive allocation by contexts of blocks. Different from traditional methods, the contexts $\boldsymbol{U}_{1,j}$ and $\boldsymbol{U}_{i,j}$ of $\boldsymbol{x}_{1,j}$ and $\boldsymbol{x}_{i,j}$ are, respectively, extracted by using the CS measurements $\boldsymbol{y}_{1,j}$ and $\boldsymbol{y}_{i,j}$, which makes CVS system more practical. After context extraction, according to the contexts $\boldsymbol{U}_{1,j}$ and $\boldsymbol{U}_{i,j}$, the numbers of CS measurements of $\boldsymbol{x}_{1,j}$ and $\boldsymbol{x}_{i,j}$ are modified as $M_{1,j}$ and $M_{i,j}$ by adaptive allocation. According to $M_{1,j}$ and $M_{i,j}$, by removing the redundancy or supplementing the insufficiency in $\boldsymbol{y}_{1,j}$ and $\boldsymbol{y}_{i,j}$, $\boldsymbol{x}_{1,j}$ and $\boldsymbol{x}_{i,j}$ are re-sampled as $\tilde{\boldsymbol{y}}_{1,j}$ and $\tilde{\boldsymbol{y}}_{i,j}$, respectively. DPCM cannot be used to quantize the adaptive measurements with different numbers. To overcome this defect of DPCM, we fuse zero padding into DPCM and predictively quantize $\tilde{\boldsymbol{y}}_{1,j}$ and $\tilde{\boldsymbol{y}}_{i,j}$ as $\tilde{\boldsymbol{y}}^{q}_{1,j}$ and $\tilde{\boldsymbol{y}}^{q}_{i,j}$. Finally, all quantized CS measurements are encoded as bits by Huffman and packaged as Packet$_k$.

Figure 3B presents the process of decoding Packet$_k$. After unpackaging Packet$_k$, the inversions of Huffman and zero-padding DPCM are implemented, and the CS measurements of $\boldsymbol{x}_{1,j}$ and $\boldsymbol{x}_{i,j}$ are recovered as $\hat{\boldsymbol{y}}_{1,j}$ and $\hat{\boldsymbol{y}}_{i,j}$ which have some quantization errors with their originals $\tilde{\boldsymbol{y}}_{1,j}$ and $\tilde{\boldsymbol{y}}_{i,j}$. The distributed reconstruction is performed to reconstruct the key frame $\boldsymbol{f}_1$ and the non-key frames $\{\boldsymbol{f}_i\}_{i=2}^{L}$. To suppress the blocking artifacts in the reconstructed frames, we realize the recovery of large blocks by merging the CS measurements

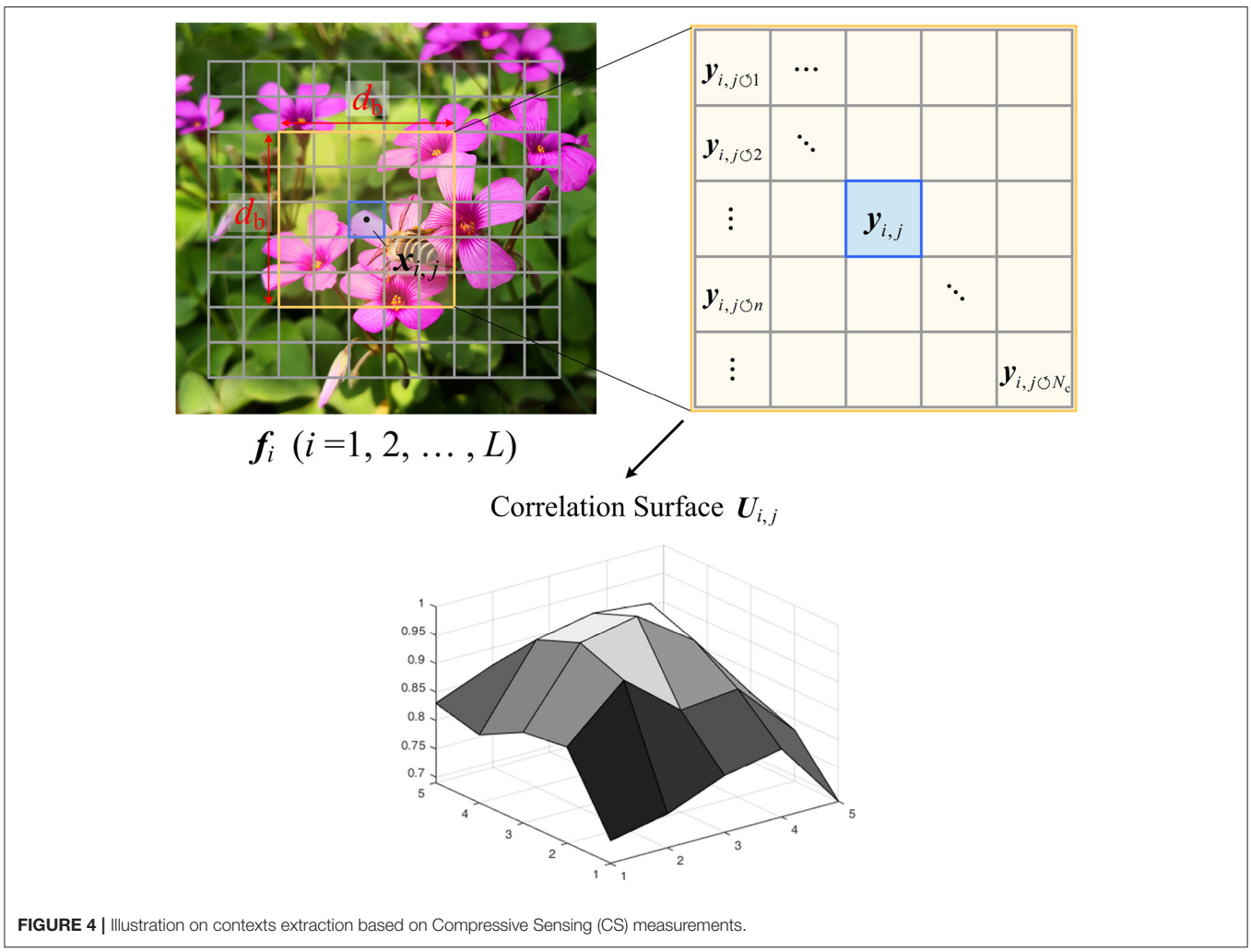


FIGURE 4 | Illustration on contexts extraction based on Compressive Sensing (CS) measurements.

of the spatially neighboring blocks, so the CS measurements of $\boldsymbol{f}_1$ and $\boldsymbol{f}_i$ are updated as $\boldsymbol{z}_{1,r}$ and $\boldsymbol{z}_{i,r}$ for large blocks. Based on $\boldsymbol{z}_{1,r}$, the reconstructed key frame $\hat{\boldsymbol{f}}_1$ is produced by using a linear recovery model, which rapidly recovers each block by a matrix-vector product. Regarding the previous and the next reconstructed key frames as references, the MH prediction outputs the reconstructed non-key frame $\hat{\boldsymbol{f}}_i$ by using $\boldsymbol{z}_{i,r}$. Finally, all reconstructed frames are combined into $\widehat{\mathrm{GOP}}_k$. Details of the core parts, including contexts extraction, measurements allocation, zero-padding DPCM, and distribution reconstruction, are described in the following subsections.

Algorithm 1 | Allocating the appropriate numbers of Compressive Sensing (CS) measurements to blocks.

Require: S_i - Subrate of $\boldsymbol{f}_i$, $P_{i,j}$ - Distribution on the sparsity of blocks $\boldsymbol{x}_{i,j}$, $j = 1, 2, \cdots, J$, N - Total number of pixels in $\boldsymbol{f}_i$, N_b - Block length;

Initial measurement number $m_{i,j}^{(0)} = \mathrm{Round}\left(P_{i,j} S_{i,j} N\right)$, and Round $(\cdot)$ is a rounding operator;
Restrict $m_{i,j}^{(0)}$ to not be larger than $0.9N_\mathrm{b}$, i.e., $m_{i,j}^{(0)} = \mathrm{Min}\left(m_{i,j}^{(0)}, 0.9N_\mathrm{b}\right)$, in which Min $(\cdot)$ is a minimization operator;
Set $M_\mathrm{sup} = S_i \cdot N - \sum_{j=1}^{J} m_{i,j}^{(0)}$, and $iter = 0$;
while $M_\mathrm{sup} > 0$, increment $iter$ by 1 **do**
  if $M_\mathrm{sup} < J$ **then**
    Randomly select M_sup blocks, and their measurement numbers are incremented by 1;
    Update $m_{i,j}^{(iter)}$, and set $M_{i,j} = m_{i,j}^{(iter)}$;
    Break;
  else
    $m_{i,j}^{(iter+1)} = m_{i,j}^{(iter)} + 1$
    $M_\mathrm{sup} = M_\mathrm{sup} - J$
  end if
end while
return $M_{i,j}, j = 1, 2, \cdots, J$.

3.2. Context Extraction

In the proposed CVS system, the context features are extracted by using the CS measurements of blocks. As illustrated in **Figure 4**, we compute the correlation surface $\boldsymbol{U}_{i,j}$ of $\boldsymbol{x}_{i,j}$ in $\boldsymbol{f}_i$ as its contexts, in which $i = 1, 2, \cdots, L$. In the surrounding window of size $d_\mathrm{b} \times d_\mathrm{b}$ centered on $\boldsymbol{x}_{i,j}$, we cannot extract the original blocks pixel-by-pixel due to the unavailability of original pixels, but can only use the CS measurements $\left\{\boldsymbol{y}_{i,j\bigcirc n}\right\}_{n=1}^{N_\mathrm{c}}$ of non-overlapping blocks $\left\{\boldsymbol{x}_{i,j\bigcirc n}\right\}_{n=1}^{N_\mathrm{c}}$, in which $N_\mathrm{c} = d_\mathrm{b}^2$. According to CS theory, the measurement matrix $\boldsymbol{\Phi}_{i,j}$ holds the Restricted Isometry Property (RIP) (Candès and Wakin, 2008) for blocks $\left\{\boldsymbol{x}_{i,j}\right\}_{j=1}^{J}$, which implies that all pairwise distances between original blocks can be well preserved in the measurement space, i.e.,

$$\begin{aligned} \|\boldsymbol{x}_{i,j} - \boldsymbol{x}_{i,j\bigcirc n}\|_2 &\approx \|\boldsymbol{\Phi}_{i,j} \cdot \boldsymbol{x}_{i,j} - \boldsymbol{\Phi}_{i,j} \cdot \boldsymbol{x}_{i,j\bigcirc n}\|_2 \\ &= \|\boldsymbol{y}_{i,j} - \boldsymbol{y}_{i,j\bigcirc n}\|_2, \forall n \in \{1, 2, \cdots, N_\mathrm{c}\} \end{aligned} \tag{10}$$

where it is noted that all blocks share the same measurement matrix $\boldsymbol{\Phi}_{i,j}$ due to the uniform allocation. Based on Equation

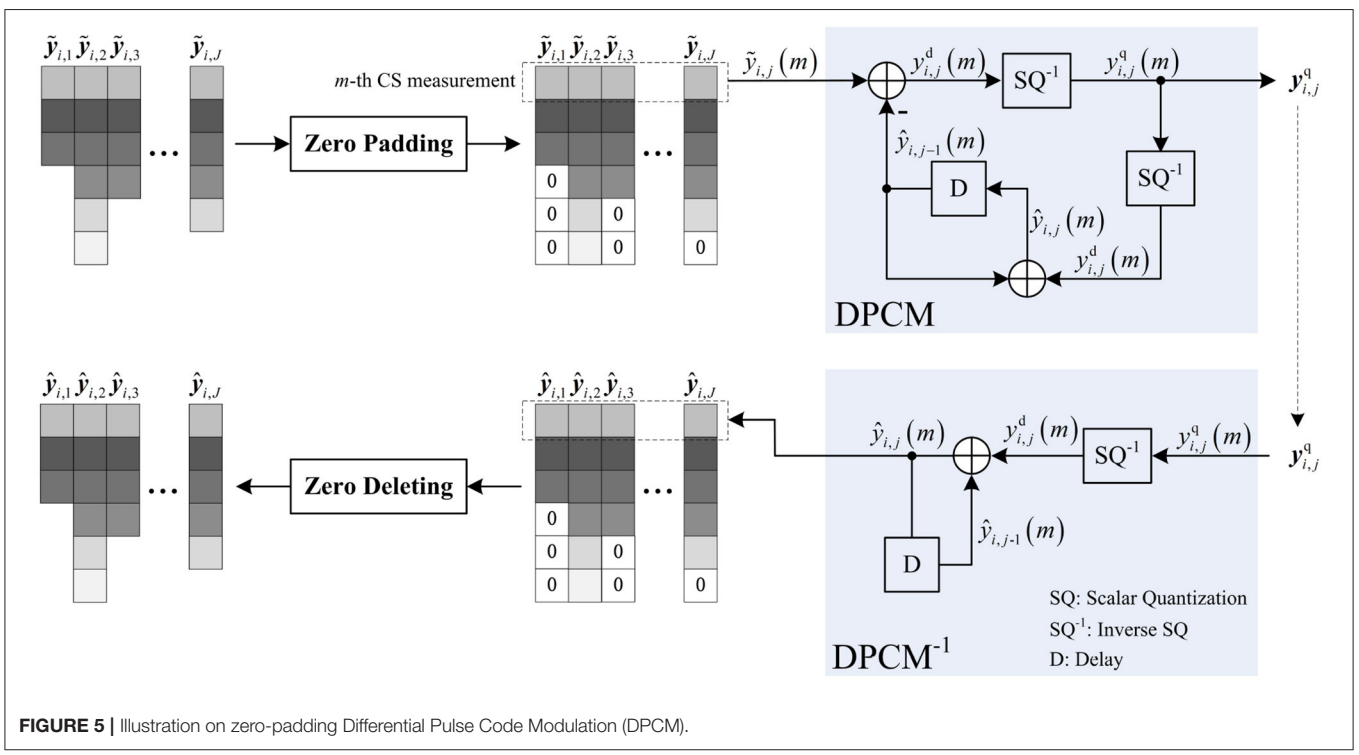


FIGURE 5 | Illustration on zero-padding Differential Pulse Code Modulation (DPCM).

(10), the similarity weights between $\boldsymbol{x}_{i,j}$ and $\boldsymbol{x}_{i,j\circlearrowleft n}$ can be estimated by

$$s_{i,j\circlearrowleft n} = \exp\left\{-\frac{\|\boldsymbol{y}_{i,j} - \boldsymbol{y}_{i,j\circlearrowleft n}\|_2^2}{2\sigma^2}\right\}, \forall n \in \{1, 2, \cdots, N_c\} \quad (11)$$

All weights constitute the correlation surface $\boldsymbol{U}_{i,j}$ as follows:

$$\boldsymbol{U}_{i,j} = \left[s_{i,j\circlearrowleft n} \,|\forall n \in \{1, 2, \cdots, N_c\}\right] \quad (12)$$

To compactly represent the contexts of $\boldsymbol{x}_{i,j}$, we compute the mean $u_{i,j}$ of $\boldsymbol{U}_{i,j}$ as the context feature, i.e.,

$$u_{i,j} = \frac{1}{N_c}\sum_{n=1}^{N_c} s_{i,j\circlearrowleft n} \quad (13)$$

3.3. Measurement Allocation

By exploiting the context feature $u_{i,j}$ of $\boldsymbol{x}_{i,j}$, we set the appropriate number of CS measurements for $\boldsymbol{x}_{i,j}$, and remove the redundancy or supplement the insufficiency in $\boldsymbol{y}_{i,j}$. The magnitudes of context features are high in smooth regions, and the magnitudes are low in the edge and texture regions, so it is found that the experience that the context feature is inversely proportional to the sparsity. Based on this experience, we can describe the distribution on the sparsity degrees of blocks by

$$P_{i,j} = \frac{u_{i,j}^{-1}}{\sum_{j=1}^{J} u_{i,j}^{-1}} \quad (14)$$

According to the present subrate S_i of $\boldsymbol{f}_i$, we construct the allocation model of CS measurements for blocks as follows:

$$\begin{aligned} M_{i,j} &= arg\min_{m_{i,j}} \sum_{j=1}^{J} \left(m_{i,j} - P_{i,j}\cdot S_i \cdot N\right) \\ \text{s.t.} &\sum_{j=1}^{j} m_{i,j} = S_i \cdot N, m_{i,j} \le 0.9 \cdot N_b, m_{i,j} \in \mathbb{N}^+ \end{aligned} \quad (15)$$

where N is the total number of pixels in $\boldsymbol{f}_i$, N_b is the block length, $m_{i,j}$ is a positive integer, and its upper bound is set to be $0.9 \cdot N_b$. The model (15) is solved according to **Algorithm 1** and outputs the final number $M_{i,j}$ of CS measurements for $\boldsymbol{x}_{i,j}$.

3.4. Zero-Padding DPCM

Due to the adaptive allocation, the lengths of the re-sampled CS measurements $\left\{\tilde{y}_{i,j}\right\}_{j=1}^{L}$ vary. Compared with SQ, DPCM provides better rate-distortion performance by adding the

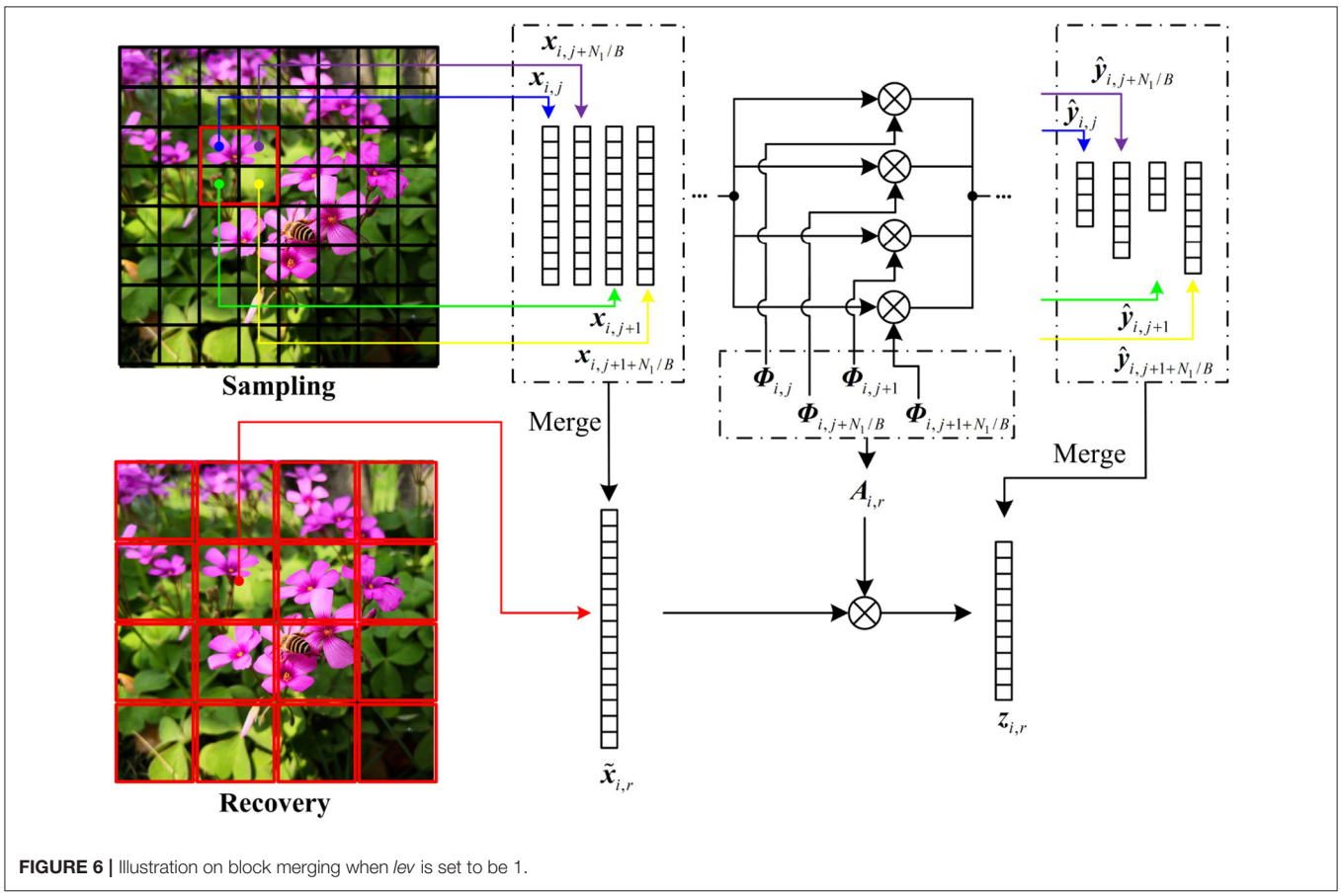


FIGURE 6 | Illustration on block merging when *lev* is set to be 1.

predictive scheme into the quantization of block-based CS measurements. However, DPCM requires that all blocks have the same number of CS measurements, as a result, DPCM cannot be used to quantize $\left\{\tilde{\boldsymbol{y}}_{i,j}\right\}_{j=1}^{L}$. To make DPCM adapt to the adaptive allocation, we propose zero-padding DPCM, whose implementation is shown in **Figure 5**. Before inputting $\tilde{\boldsymbol{y}}_{i,j}$ to DCPM, we fill zeros in the last of $\tilde{\boldsymbol{y}}_{i,j}$ to make its length the same as others. After obtaining the de-quantized CS measurements $\hat{\boldsymbol{y}}_{i,j}$, we delete the zeros in the last of $\hat{\boldsymbol{y}}_{i,j}$ to recover its original length $M_{i,j}$. By zero padding, each measurement in $\hat{\boldsymbol{y}}_{i,j-1}$ can be used to predict the corresponding measurement in $\hat{\boldsymbol{y}}_{i,j}$, and especially when there is predictive measurement $\hat{y}_{i,j-1}(m)$ of the m-th measurement $\tilde{y}_{i,j}(m)$, the residual $y_{i,j}^{\mathrm{d}}(m)$ can be significantly reduced due to the intrinsic spatial correlation between $\tilde{\boldsymbol{y}}_{i,j}$ and $\tilde{\boldsymbol{y}}_{i,j-1}$. The rate-distortion curves of the reconstructed *Foreman*, *Mobile*, and *Football* sequences are presented when zero-padding DPCM and SQ are, respectively, used to quantize the adaptive CS measurements (shown in **Supplementary Figure 1**), in which the rate-distortion curve is measured in terms of the Peak Signal-to-Noise Ratio (PSNR) in dB and bitrate in bits per pixel (bpp), and the linear recovery algorithm presented in subsection 3.5 is used to recover each video frame. It can be seen that zero-padding DPCM presents competitive performance with SQ at low bitrates but as the bitrate increases, its improvement of performance over SQ is increasingly significant. From these results, we find that the efficiency of zero-padding DPCM relies on the correlation between block-based CS measurements. With insufficient measurements, the correlation is weakened by the filling of excessive zeros, causing the performance degradation, but when measurements are sufficient, a high correlation is maintained, so the performance improvement stands out. From the above, zero-padding DPCM is more suitable for adaptive measurements compared with SQ.

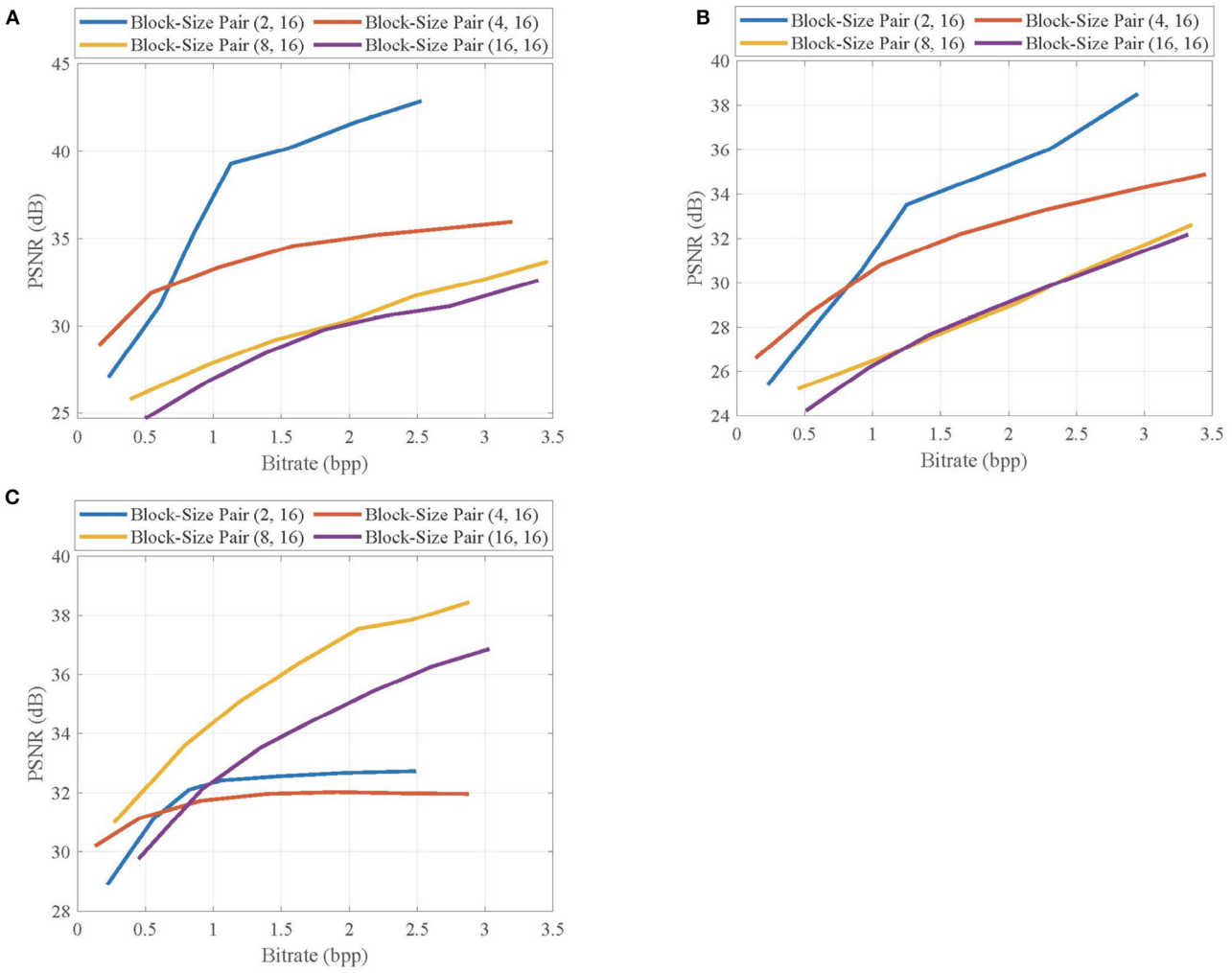


FIGURE 7 | Rate-distortion curves of the reconstructed key frames in **(A)** *Foreman*, **(B)** *BlowingBubbles*, and **(C)** *ParkScene* sequences at different block-size pairs.

3.5. Distributed Reconstruction

At decoder, the distributed strategy is performed to reconstruct the key frame $\boldsymbol{f}_1$ and the non-key frames $\{\boldsymbol{f}_i\}_{i=2}^{L}$, in which $\boldsymbol{f}_1$ is estimated by a linear recovery model, and $\boldsymbol{f}_i$ is produced by MH prediction. To highlight the complex structures by contexts, a small block size is more desired at the encoder. However, the small block size causes serious blocking artifacts due to the differences of neighboring blocks in recovery quality. To suppress the blocking artifacts, we merge the CS measurements $\{\hat{\boldsymbol{y}}_{i,j}\}_{j=1}^{J}$ of the small blocks $\{\boldsymbol{x}_{i,j}\}_{j=1}^{J}$ into those $\{\boldsymbol{z}_{i,r}\}_{r=1}^{R}$ of the large blocks $\{\tilde{\boldsymbol{x}}_{i,r}\}_{r=1}^{R}$ and realize the sampling of small blocks and the recovery of large blocks. The size $B_{lev} \times B_{lev}$ of large block is set to be

$$B_{lev} = 2^{lev} \cdot B, lev = 1, 2, \cdots \tag{16}$$

in which lev is a positive integer. The number R of large blocks is N/B_{lev}^2, and it is smaller than the number J of small blocks. **Figure 6** illustrates the block merging when lev is set to be 1. The four neighboring blocks $\boldsymbol{x}_{i,j}$, $\boldsymbol{x}_{i,j+1}$, $\boldsymbol{x}_{i,j+N_1/B}$, $\boldsymbol{x}_{i,j+1+N_1/B}$ are merged into a large block $\tilde{\boldsymbol{x}}_{i,r}$, and their CS measurements $\hat{\boldsymbol{y}}_{i,j}$, $\hat{\boldsymbol{y}}_{i,j+1}$, $\hat{\boldsymbol{y}}_{i,j+N_1/B}$, and $\hat{\boldsymbol{y}}_{i,j+1+N_1/B}$ are spliced into $\boldsymbol{z}_{i,r}$ in rows, i.e.,

$$\boldsymbol{z}_{i,r} = \begin{bmatrix} \hat{\boldsymbol{y}}_{i,j} \\ \hat{\boldsymbol{y}}_{i,j+1} \\ \hat{\boldsymbol{y}}_{i,j+N_1/B} \\ \hat{\boldsymbol{y}}_{i,j+1+N_1/B} \end{bmatrix} \approx \boldsymbol{\Lambda}_{i,r} \cdot \begin{bmatrix} \boldsymbol{x}_{i,j} \\ \boldsymbol{x}_{i,j+1} \\ \boldsymbol{x}_{i,j+N_1/B} \\ \boldsymbol{x}_{i,j+1+N_1/B} \end{bmatrix} \tag{17}$$

$$\boldsymbol{\Lambda}_{i,r} = \begin{bmatrix} \boldsymbol{\Phi}_{i,j} & & & 0 \\ & \boldsymbol{\Phi}_{i,j+1} & & \\ & & \boldsymbol{\Phi}_{i,j+N_1/B} & \\ 0 & & & \boldsymbol{\Phi}_{i,j+1+N_1/B} \end{bmatrix} \tag{18}$$

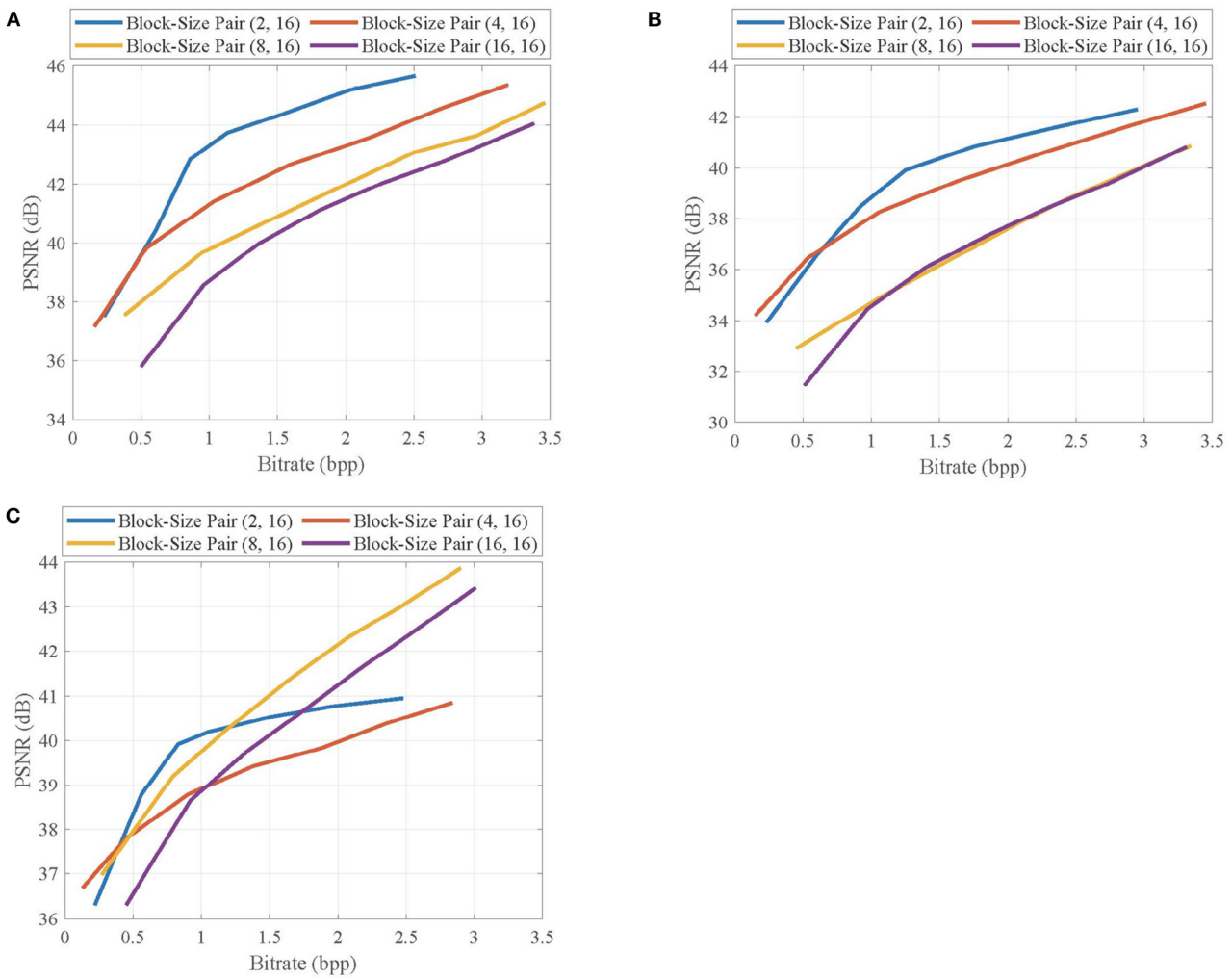


FIGURE 8 | Rate-distortion curves of the reconstructed non-key frames in **(A)** *Foreman*, **(B)** *BlowingBubbles*, and **(C)** *ParkScene* sequences at different block-size pairs.

in which $\boldsymbol{\Lambda}_{i,r}$ is the diagonal matrix composed of the block measurement matrices $\boldsymbol{\Phi}_{i,j}$, $\boldsymbol{\Phi}_{i,j+1}$, $\boldsymbol{\Phi}_{i,j+N_1/B}$, and $\boldsymbol{\Phi}_{i,j+1+N_1/B}$, N_1 is the total number of rows in $\boldsymbol{f}_i$, and B is the block size of the small block. To make $\boldsymbol{z}_{i,r}$, the CS measurements of $\tilde{\boldsymbol{x}}_{i,r}$, we transform $\tilde{\boldsymbol{x}}_{i,r}$ as

$$\begin{bmatrix} \boldsymbol{x}_{i,j} \\ \boldsymbol{x}_{i,j+1} \\ \boldsymbol{x}_{i,j+N_1/B} \\ \boldsymbol{x}_{i,j+1+N_1/B} \end{bmatrix} = \boldsymbol{I} \cdot \tilde{\boldsymbol{x}}_{i,r} \tag{19}$$

in which $\boldsymbol{I}$ is an elementary column transformation matrix. Plugging Equation (19) into Equation (17), we build the bridge between $\tilde{\boldsymbol{x}}_{i,r}$ and $\boldsymbol{z}_{i,r}$ by

$$\boldsymbol{z}_{i,r} \approx \boldsymbol{\Lambda}_{i,r} \cdot \boldsymbol{I} \cdot \tilde{\boldsymbol{x}}_{i,r} = \boldsymbol{A}_{i,r} \cdot \tilde{\boldsymbol{x}}_{i,r} \tag{20}$$

in which $\boldsymbol{A}_{i,r} = \boldsymbol{\Lambda}_{i,r} \cdot \boldsymbol{I}$. According to Equation (20), the large block $\tilde{\boldsymbol{x}}_{i,r}$ can be recovered by using $\boldsymbol{z}_{i,r}$. When *lev* is set to be larger than 1, the block merging can be done in manner similar to the above.

After the block merging, we use $\{\boldsymbol{z}_{1,r}\}_{r=1}^{R}$ to recover the key frame $\boldsymbol{f}_1$. The block $\tilde{\boldsymbol{x}}_{1,r}$ of $\boldsymbol{f}_1$ is linearly estimated by

$$\hat{\boldsymbol{x}}_{1,r} = \boldsymbol{P}_{1,r} \cdot \boldsymbol{z}_{1,r} \tag{21}$$

in which $\boldsymbol{P}_{1,r}$ is the transformation matrix produced by the following model:

$$\boldsymbol{P}_{1,r} = arg \min_{\boldsymbol{P}} \left\{ E\left[\| \tilde{\boldsymbol{x}}_{1,r} - \boldsymbol{P} \cdot \boldsymbol{z}_{1,r} \|_2^2 \right] \right\} \tag{22}$$

in which $E[\cdot]$ denotes the expectation function. The model (22) outputs the optimal transformation matrix to minimize the mean square error between $\tilde{\boldsymbol{x}}_{1,r}$ and its estimator $\hat{\boldsymbol{x}}_{1,r}$, and it can be

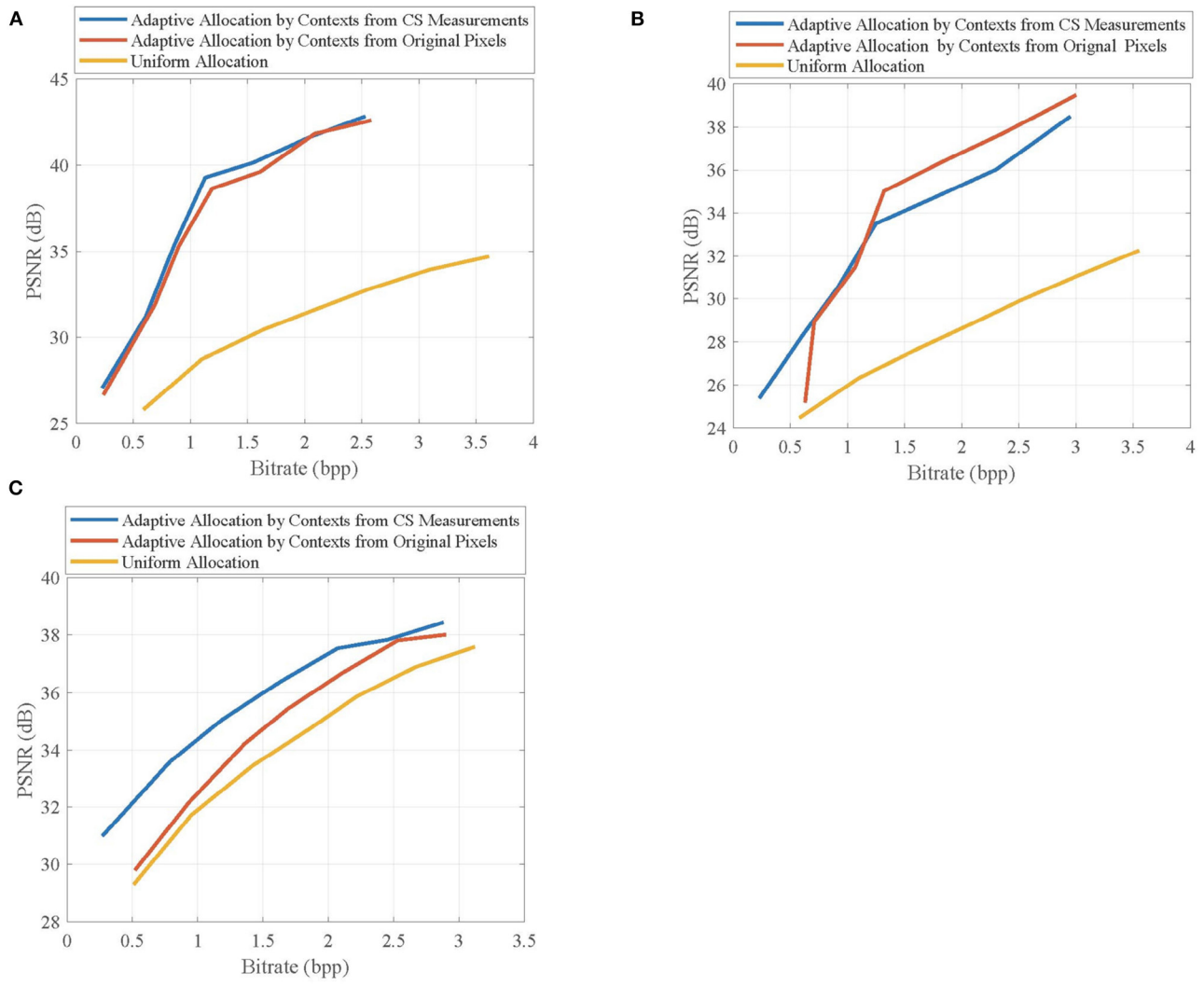


FIGURE 9 | Rate-distortion curves of the reconstructed key frames in **(A)** *Foreman*, **(B)** *BlowingBubbles*, and **(C)** *ParkScene* sequences when using different allocation schemes. For *Foreman* and *BlowingBubbles*, the block-size pair is set to be (2, 16), and for *ParkScene*, the block-size is set to be (8, 16).

solved by making the gradient of objective function equal to 0, producing

$$\boldsymbol{P}_{1,r} = E\left[\tilde{\boldsymbol{x}}_{1,r}\boldsymbol{z}_{1,r}^{\mathrm{T}}\right] E^{-1}\left[\boldsymbol{z}_{1,r}\boldsymbol{z}_{1,r}^{\mathrm{T}}\right] \tag{23}$$

Plugging Equation (20) into Equation (23), we get

$$\boldsymbol{P}_{1,r} = \boldsymbol{Cor}_{\mathrm{xx}} \cdot \boldsymbol{A}_{1,r}^{\mathrm{T}} \left(\boldsymbol{A}_{1,r} \cdot \boldsymbol{Cor}_{\mathrm{xx}} \cdot \boldsymbol{A}_{1,r}^{\mathrm{T}}\right)^{-1} \tag{24}$$

$$\boldsymbol{Cor}_{\mathrm{xx}} = E\left[\tilde{\boldsymbol{x}}_{1,r}\tilde{\boldsymbol{x}}_{1,r}^{\mathrm{T}}\right] \tag{25}$$

in which $\boldsymbol{Cor}_{\mathrm{xx}}$ is the auto-correlation matrix of $\tilde{\boldsymbol{x}}_{1,r}$, and its element $Cor_{\mathrm{xx}}[m,n]$ is estimated as follows:

$$Cor_{\mathrm{xx}}[m,n] = 0.95^{\delta_{m,n}} \tag{26}$$

in which $\delta_{m,n}$ is the Euclidean distance between two pixels $\tilde{x}_{1,r}(m)$ and $\tilde{x}_{1,r}(n)$ in $\tilde{\boldsymbol{x}}_{1,r}$. When the subrate is set to be large, the linear recovery model can provide excellent visual quality while costing fewer computations.

4. EXPERIMENTAL RESULTS

We evaluate the proposed CVS system on video sequences with various resolutions, including seven CIF (352 × 288) sequences *Akiyo*, *Bus*, *Container*, *Coastguard*, *Football*, *Foreman*, *Hall*, one WQVGA (416 × 240) sequence *BlowingBubbles*, and one 1080p (1920 × 1080) sequence *ParkScene*. In the proposed CVS system, the window size $d_{\mathrm{b}} \times d_{\mathrm{b}}$ and the normalization factor σ are, respectively, set to be 11 × 11 and 10 for the context extraction, the window size $W \times W$ and the regularization factor β are, respectively, set to be 21 × 21 and 0.25 for the MH prediction, and

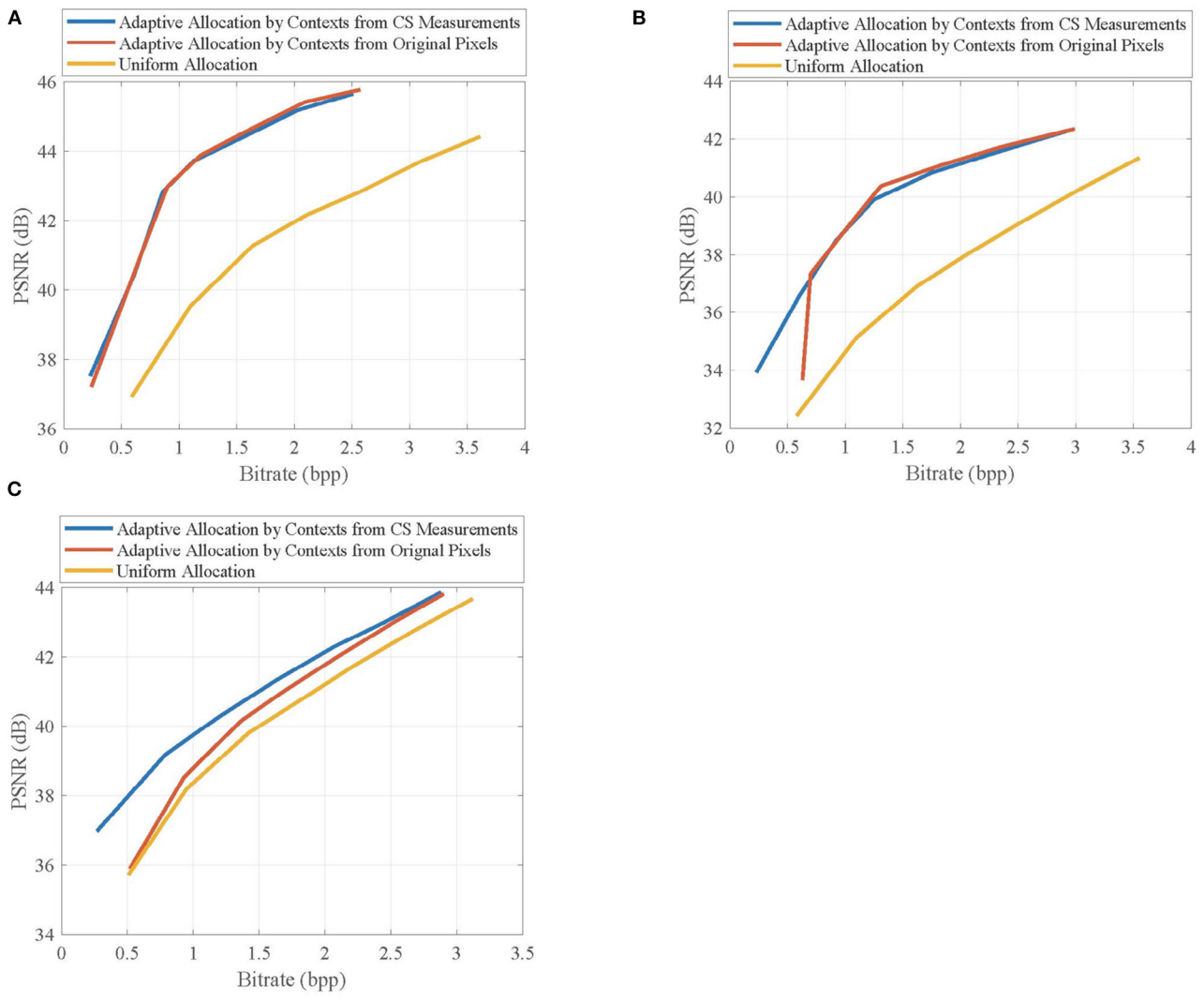


FIGURE 10 | Rate-distortion curves of the reconstructed non-key frames in **(A)** *Foreman*, **(B)** *BlowingBubbles*, and **(C)** *ParkScene* sequences when using different allocation schemes. For *Foreman* and *BlowingBubbles*, the block-size pair is set to be (2, 16), and for *ParkScene*, the block-size is set to be (8, 16).

the measurement matrix is produced by Gaussian distribution. First, we discuss the effects of different block sizes on the proposed CVS system. Second, we evaluate the performance improvement resulting from the used context extraction. Finally, we compare the proposed CVS system with two state-of-the-art CVS systems: SS-CVS (Li et al., 2020) and MH-RTIK (Chen C. et al., 2020) in terms of the rate-distortion performance. PSNR is used to evaluate the qualities of reconstructed video sequences, and the bitrate denotes the average amount of bits per pixel to encode a video sequence. The variation of PSNR with bitrate is called the rate-distortion performance. The computational complexity is measured by the execution time. Experiments are implemented with MATLAB on a workstation with 3.30-GHz CPU and 8 GB RAM.

4.1. Effects of Block Sizes

In the proposed CVS system, in order to highlight the complex structures by contexts, we desire a small block size at encoder, but at decoder, a large block size is desired to suppress the blocking artifacts in the reconstructed video frames. We set a block-size pair (B, B_{lev}), in which B and B_{lev} are the block sizes for sampling and recovery, respectively, and evaluate the effects of different block-size pairs on the reconstruction qualities of key frames and non-key frames.

First, we select the first frames of *Foreman*, *BlowingBubbles*, and *ParkScene* sequences as the key frames, which are linearly recovered, and show their rate-distortion curves at different block-size pairs in **Figure 7**. For *Foreman* and *BlowingBubbles* with the low resolution, the block-size pair (4, 16) achieves higher PSNR values than others with low bitrates, but the rate-distortion curve for the block-size pair (2, 16) rapidly increases as the bitrate increases and significant PSNR gains are achieved when compared with other block-size pairs. These results indicate that the small blocks used in adaptive allocation and large blocks for linear recovery fit together well. For *ParkScene* with high resolution, when the block size B for sampling is set to be too small, e.g., $B = 2$, no block can contain sufficient structures, causing the rate-distortion performance to degenerate as the bitrate increases, but a suitable block size for sampling is set, e.g., $B = 8$, PSNR gains can be significantly improved.

Then, we select the second frames of *Foreman*, *BlowingBubbles*, and *ParkScene* sequences as the non-key frames, which are recovered by MH prediction based on the reconstructed previous and next key frames at the subrate 0.7, and show their rate-distortion curves at different block-size pairs in **Figure 8**. Similar to the results from key frames, for *Foreman* and *BlowingBubbles*, the better rate-distortion performance is achieved when the block-size pair is set to be (2, 16), and for *ParkScene*, in order to prevent the loss of structures, the block size for sampling is appropriately set to be 8.

Given the above, we can see that the bad effects resulting from the extraction of contexts can be suppressed by the block

TABLE 1 | Average Peak Signal-to-Noise Ratio (PSNR) (dB) for reconstructed video sequences by the proposed Compressive Video Sensing (CVS) system, Scalable Structured CVS (SS-CVS) (Trevisi et al., 2020), and Multi-Hypothesis Reweighted TIKhonov (MH-RTIK) (Chen C. et al., 2020) at subrates 0.1 to 0.5.

Sequence	Resolution	Algorithm	Subrate S_{NK}				
			0.1	0.2	0.3	0.4	0.5
GOP Length $L = 2$							
Container	CIF	MH-RTIK	33.67	34.76	35.08	35.28	35.47
		Proposed	**38.74**	**39.92**	**40.38**	**40.47**	**40.61**
Coastguard		MH-RTIK	33.12	34.26	34.69	35.08	35.43
		Proposed	**35.80**	**37.22**	**38.30**	**38.89**	**39.45**
Hall		MH-RTIK	37.10	38.01	38.39	38.65	38.91
		Proposed	**38.26**	**39.69**	**40.82**	**41.23**	**41.50**
Foreman		MH-RTIK	36.52	37.09	37.56	37.96	38.60
		Proposed	**38.13**	**39.66**	**40.87**	**41.41**	**41.78**
GOP Length $L = 10$							
Akiyo	CIF	SS-CVS	17.70	24.80	33.06	36.55	39.23
		Proposed	**40.75**	**43.50**	**45.28**	**45.09**	**45.56**
Bus		SS-CVS	18.65	23.57	25.71	27.67	30.10
		Proposed	**25.65**	**38.31**	**30.97**	**32.97**	**34.00**
Football		SS-CVS	15.52	23.95	27.87	30.33	32.93
		Proposed	**28.98**	**32.67**	**35.78**	**36.55**	**37.28**
Foreman		SS-CVS	13.40	20.51	28.07	32.90	35.25
		Proposed	**33.18**	**36.00**	**38.55**	**39.54**	**40.22**
BlowingBubble	QWVGA	SS-CVS	16.93	23.50	28.47	30.70	32.84
		Proposed	**30.13**	**32.17**	**33.58**	**35.01**	**35.68**
ParkScene	1080P	SS-CVS	23.19	30.04	33.14	35.53	36.62
		Proposed	**33.01**	**35.18**	**36.79**	**37.97**	**38.67**

merging, therefore, the quality improvement from contexts-based allocation is further enhanced.

4.2. Effects of Contexts

In the proposed CVS system, the contexts are extracted from CS measurements and used to adaptively allocate the CS measurements for blocks, leading to the improvement of reconstruction quality. To verify the validity of contexts from CS measurements on the quality improvement, we evaluate the effects of different allocation schemes on the rate-distortion performance of the proposed CVS system. The uniform allocation is used as a benchmark, and the adaptive allocation uses the contexts extracted from CS measurements and original pixels, respectively.

Figure 9 shows the rate-distortion curves of the reconstructed key frames when using different allocation schemes, in which the key frames are, respectively, taken from the first frames of *Foreman*, *BlowingBubbles*, and *ParkScene* sequences. It can be seen that adaptive allocation outperforms uniform allocation in PSNR values at any bitrate, indicating that contexts contribute to quality improvement. Importantly, the contexts from CS measurements are competitive with those from original pixels, and their performance gaps are very small, which means that CS measurements can better represent the contexts of blocks.

Figure 10 shows the rate-distortion curves of the reconstructed non-key frames when using different allocation schemes, in which the non-key frames are, respectively, taken from the second frames of *Foreman*, *BlowingBubbles*, and *ParkScene* sequences. It can be seen that the adaptive allocation is still effective for MH prediction, and it can significantly improve the rate-distortion performances when compared with uniform allocation. The contexts from CS measurements have similar efficiency of allocation to that of contexts from original pixels, which proves that the merits of adaptive allocation can still be maintained in the measurement domain.

The above results indicate that the contexts extracted by CS measurements prompt the adaptive allocation to improve the reconstruction quality of CVS system, which makes the

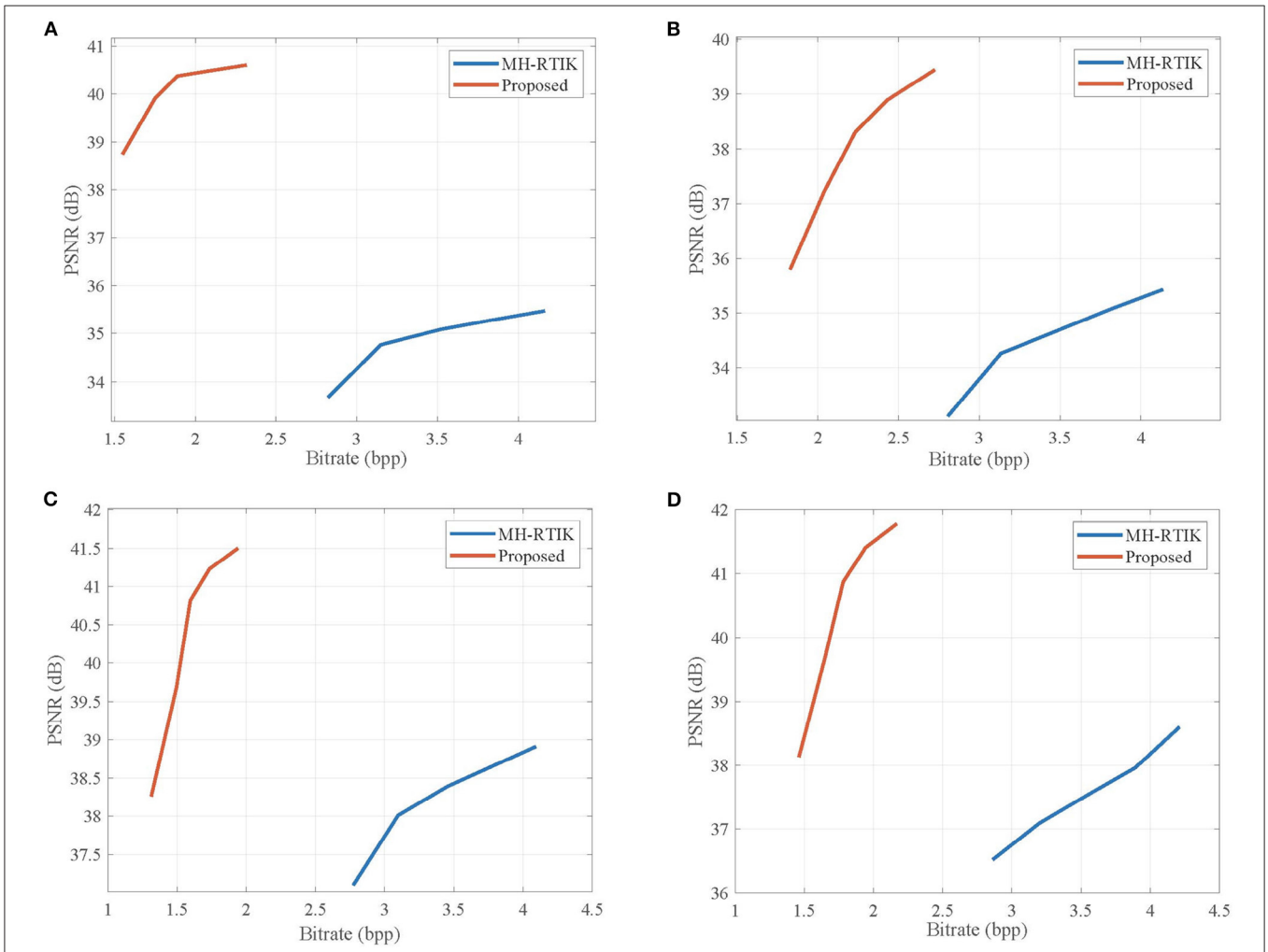


FIGURE 11 | Rate-distortion curves obtained by the proposed CVS system and Multi-Hypothesis TIKhonov (MH-TIK) (Chen C. et al., 2020) for **(A)** *Container*, **(B)** *Coastguard*, **(C)** *Hall*, and **(D)** *Foreman* sequences. Note that the length *L* of GOP is set to be 2.

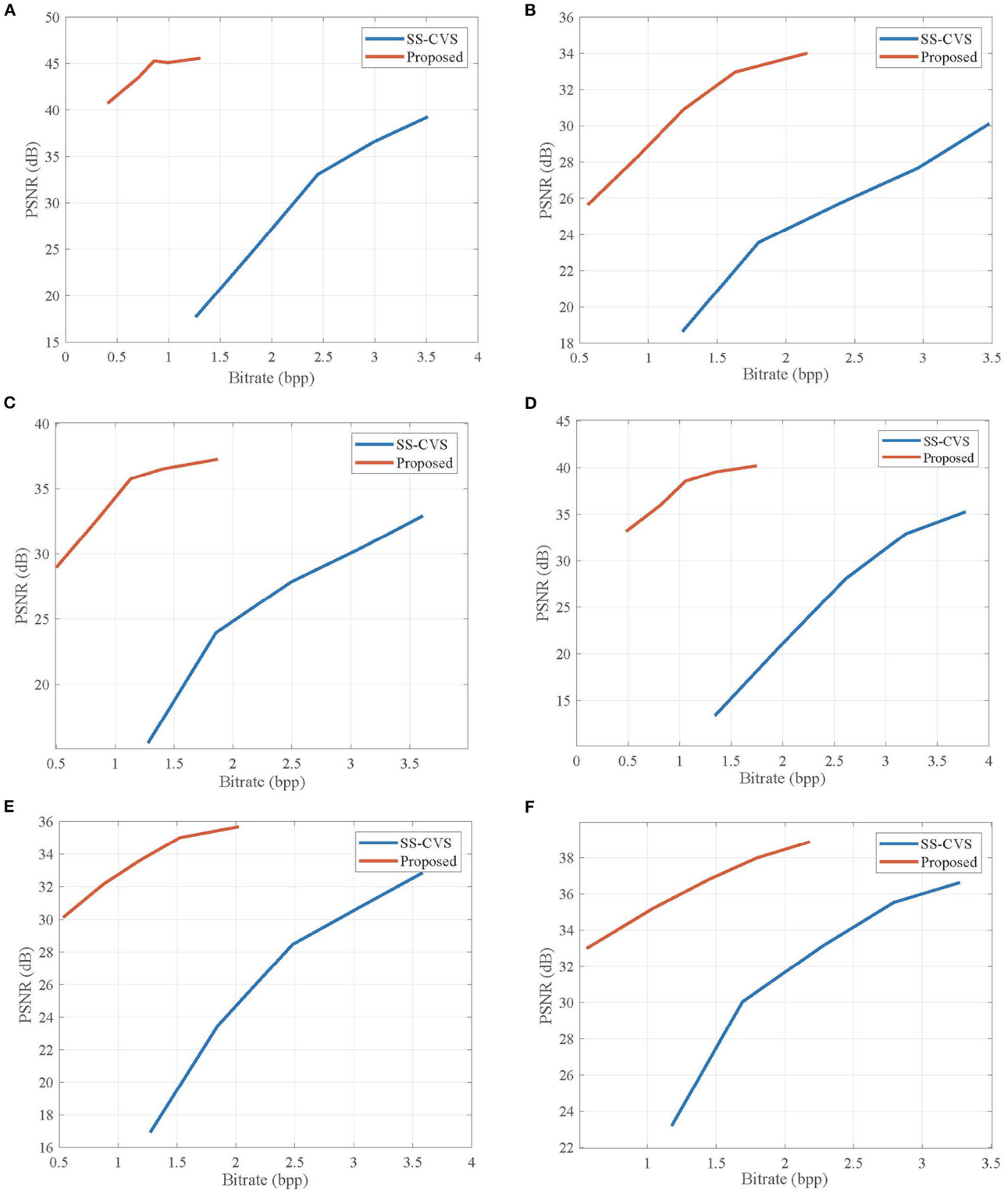


FIGURE 12 | Rate-distortion curves obtained by the proposed CVS system and Scalable Structured CVS (SS-CVS) (Li et al., 2020) for **(A)** *Akiyo*, **(B)** *Bus*, **(C)** *Football*, **(D)** *Foreman*, **(E)** *BlowingBubble*, and **(F)** *ParkScene* sequences. Note that the length *L* of GOP is set to be 10.

TABLE 2 | Average encoding time (s/frame) and decoding time (s/frame) on video sequences with different resolutions for the proposed CVS system, SS-CVS (Li et al., 2020), and MH-RTIK (Chen C. et al., 2020).

Resolution	Algorithm	Encoding Time (s/frame)	Decoding Time (s/frame)
Average on Subrates S_{NK} 0.1 to 0.5			
CIF	MH-RTIK	0.17	19.34
	Proposed	0.63	4.48
Average on Subrates $S_{NK} = 0.6$			
CIF	SS-CVS	5.40	21.22
	Proposed	0.64	7.79
QWVGA	SS-CVS	4.90	17.23
	Proposed	0.64	7.69
1080P	SS-CVS	108.10	401.8
	Proposed	1.83	162.47

proposed CVS system more suitable to the applications with limited resources.

4.3. Performance Comparisons

We evaluate the performance of the proposed CVS system by comparing it with the two state-of-the-art CVS systems: SS-CVS (Li et al., 2020) and MH-RTIK (Chen C. et al., 2020). To make a fair comparison, we keep the parameter settings of SS-CVS and MH-RTIK in their original reports, some important details are repeated as follows:

1) SS-CVS: the system consists of one base layer and one enhancement layer; the block size is set to be 16; the length of GOP is 10; the subrate of key frame is set to be 0.9; the dimension of the subspace is 10; the number of subspaces is 50.
2) MH-RTIK: the sub-block extraction is used; the number of hypotheses is 40; the block size is set to be 16; the length of GOP is 2; the subrate of key frame is set to be 0.7.

In addition, we employ SQ and Huffman in SS-CVS and MH-RTIK to compress the CS measurements. For the proposed CVS system, the block-size pair is set to be $(2, 16)$ for CIF and QWVGA sequences and $(8, 16)$ for 1080P sequences, the subrate S_K of key frame is set to be 0.7, the results under the GOP length $L = 2$ are compared with those of MH-RTIK, and the results under the GOP length $L = 10$ are compared with those of SS-CVS.

Table 1 lists the average PSNR values for the reconstructed video sequences by the proposed CVS system, SS-CVS, and MH-RTIK when the subrate S_{NK} of non-key frame varies from 0.1 to 0.5. Compared with MH-RTIK, the proposed CVS system achieves obvious PSNR gains at any subrate, e.g., the average PSNR gain is 2.824 dB for the *Foreman* sequence. Compared with SS-CVS, the proposed CVS system also presents higher PSNR values at any subrate, and especially for low subrates, PSNR gains are significant, e.g., when the subrate is 0.1, PSNR gains are 9.82, 13.20, and 19.78 dB for *ParkScene*, *BlowingBubble*, *Foreman* sequences, respectively. **Figures 11**, **12** show the rate-distortion curves for the proposed CVS system, MH-RTIK, and SS-CVS. Due to the implementation of zero-padding DPCM, the performance improvement of the proposed CVS system is further enhanced when compared with MH-RTIK and SS-CVS. By the objective evaluation of the reconstruction quality, it can be indicated that the proposed CVS system can significantly improve the qualities of the reconstructed video sequences.

Table 2 lists the average encoding time (s/frame) and decoding time (s/frame) on video sequences with different resolutions for the proposed CVS system, SS-CVS, and MH-RTIK. We compute the average execution time on the range $[0.1, 0.5]$ of subrate S_{NK} for the proposed CVS system and compare it with that of MH-RTIK for CIF sequences. The encoding speed of the proposed CVS system is slowed down due to the contexts-based adaptive allocation, and its encoding time is 0.63 s per frame, larger than that of MH-RTIK. Assisted by the simple linear recovery, the proposed CVS system reduces the decoding complexity, and only costs 4.48 s to reconstruct a video frame, however, MH-RTIK requires 19.34 s per frame. Under the subrate $S_{NK} = 0.6$, the execution time of the proposed CVS algorithm is compared with that of SS-CVS for the CIF, QWVGA, and 1080P video sequences, respectively. Compared with SS-CVS, the proposed CVS system costs less encoding time, and the encoding time does not dramatically increase as the resolution increases, e.g., for 1080P sequence, the proposed CVS system only costs 1.83 s per frame, but SS-CVS costs 108.10 s. In SS-CVS, the subspace clustering and the basis derivation are implemented at the encoder, and they lead to more encoding costs than the adaptive allocation in the proposed CVS system. The proposed CVS system costs less decoding time than SS-CVS, and its decoding costs also grow more slowly when compared with SS-CVS, e.g., for 1080P sequence, the proposed CVS system costs 162.47 s per frame, and the SS-CVS costs 401.8 s. The heavy computational burdens for SS-CVS derive from the non-linear subspace learning, but the decoding complexity of the proposed CVS system is limited benefiting from the linear recovery and prediction. From the above, we can see that the proposed CVS system still keeps a low computational complexity while providing better rate-distortion performance.

5. CONCLUSION

In this article, a context-based CVS system is proposed to improve the visual quality of the reconstructed video sequences. At the encoder, the CS measurements are adaptively allocated for blocks according to the contexts of video frames. Innovatively, the contexts are extracted by CS measurements. Although the extraction of contexts is independent of original pixels, these contexts can still better reveal the structural complexity of each block. To guarantee better rate-distortion performance, the zero-padding DPCM is proposed to quantize these adaptive measurements. At the decoder, the key frames are reconstructed by linear recovery, and these non-key frames are reconstructed by MH prediction. Thanks to the effectiveness of context-based adaptive allocation, the simple recovery schemes also provide the comfortable visual quality. Experimental results show that the proposed CVS system improves the rate-distortion performances

when compared with two state-of-the-art CVS systems, including MH-RTIK and SS-CVS, and guarantees a low computational complexity.

As the research in this article is exploratory, there are many intriguing questions that future work should consider. First, the estimation of block sparsity should be analyzed in mathematics. Second, we will investigate how to fuse the quantization into adaptive allocation. More importantly, we will deploy the adaptive CVS system on an actual hardware platform.

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

RL: designed the study and drafted the manuscript. YY: conducted experiments and analyzed the data. FS: critically reviewed and improved the manuscript. All authors have read and approved the final version of the manuscript.

FUNDING

This work was supported in part by the Project of Science and Technology, Department of Henan Province in China (212102210106), National Natural Science Foundation of China (31872704), Innovation Team Support Plan of University Science and Technology of Henan Province in China (19IRTSTHN014), and Guangxi Key Laboratory of Wireless Wideband Communication and Signal Processing of China.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fpls.2022.849606/full#supplementary-material

REFERENCES

Akila, I., Sivakumar, A., and Swaminathan, S. (2017). "Automation in plant growth monitoring using high-precision image classification and virtual height measurement techniques," in *2017 International Conference on Innovations in Information, Embedded and Communication Systems (ICIIECS)* (Coimbatore), 1–4.

Azghani, M., Karimi, M., and Marvasti, F. (2016). Multihypothesis compressed video sensing technique. *IEEE Trans. Circuits Syst. Video Technol.* 26, 627–635. doi: 10.1109/TCSVT.2015.2418586

Baraniuk, R. G. (2007). Compressive sensing. *IEEE Signal Process. Mag.* 24, 118–121. doi: 10.1109/MSP.2007.4286571

Baraniuk, R. G., Goldstein, T., Sankaranarayanan, A. C., Studer, C., Veeraraghavan, A., and Wakin, M. B. (2017). Compressive video sensing: algorithms, architectures, and applications. *IEEE Signal Process. Mag.* 34, 52–66. doi: 10.1109/MSP.2016.2602099

Becker S., and Bobin J., C. E. (2011). Nesta: a fast and accurate first-order method for sparse recovery. *SIAM J. Imag. Sci.* 4, 1–39. doi: 10.1137/090756855

Bigot, J., Boyer, C., and Weiss, P. (2016). An analysis of block sampling strategies in compressed sensing. *IEEE Trans. Inf. Theory* 62, 2125–2139. doi: 10.1109/TIT.2016.2524628

Candès, E. J., and Wakin, M. B. (2008). An introduction to compressive sampling. *IEEE Signal Process. Mag.* 25, 21–30. doi: 10.1109/MSP.2007.914731

Chen, C., Tramel, E. W., and Fowler, J. E. (2011). "Compressed-sensing recovery of images and video using multihypothesis predictions," in *2011 Conference Record of the Forty Fifth Asilomar Conference on Signals, Systems and Computers (ASILOMAR)* (Pacific Grove, CA), 1193–1198.

Chen, C., Zhou, C., Liu, P., and Zhang, D. (2020). Iterative reweighted tikhonov-regularized multihypothesis prediction scheme for distributed compressive video sensing. *IEEE Trans. Circuits Syst. Video Technol.* 30, 1–10. doi: 10.1109/TCSVT.2018.2886310

Chen, Y., Huang, T.-Z., He, W., Yokoya, N., and Zhao, X.-L. (2020). Hyperspectral image compressive sensing reconstruction using subspace-based nonlocal tensor ring decomposition. *IEEE Trans. Image Process.* 29, 6813–6828. doi: 10.1109/TIP.2020.2994411

Deng, C., Zhang, Y., Mao, Y., Fan, J., Suo, J., Zhang, Z., et al. (2021). Sinusoidal sampling enhanced compressive camera for high speed imaging. *IEEE Trans. Pattern Anal. Mach. Intell.* 43, 1380–1393. doi: 10.1109/TPAMI.2019.2946567

Do, T. T., Chen, Y., Nguyen, D. T., Nguyen, N., Gan, L., and Tran, T. D. (2009). "Distributed compressed video sensing," in *2009 16th IEEE International Conference on Image Processing (ICIP)* (Baltimore, MD), 1393–1396.

Do, T. T., Gan, L., Nguyen, N. H., and Tran, T. D. (2012). Fast and efficient compressive sensing using structurally random matrices. *IEEE Trans. Signal Process.* 60, 139–154. doi: 10.1109/TSP.2011.2170977

Gan, L. (2007). "Block compressed sensing of natural images," in *2007 15th International Conference on Digital Signal Processing* (Cardiff), 403–406.

Gao, X., Zhang, J., Che, W., Fan, X., and Zhao, D. (2015). "Block-based compressive sensing coding of natural images by local structural measurement matrix," in *2015 Data Compression Conference* (Snowbird, UT), 133–142.

Girod, B., Aaron, A., Rane, S., and Rebollo-Monedero, D. (2005). Distributed video coding. *Proc. IEEE* 93, 71–83. doi: 10.1109/JPROC.2004.839619

Grimblatt, V., Jégo, C., Ferré, G., and Rivet, F. (2021). How to feed a growing population—an iot approach to crop health and growth. *IEEE J. Emerg. Sel. Top. Circuits Syst.* 11, 435–448. doi: 10.1109/JETCAS.2021.3099778

Guo, L., Liu, Y., Hao, H., Han, J., and Liao, T. (2018). "Growth monitoring and planting decision supporting for pear during the whole growth stage based on pie-landscape system," in *2018 7th International Conference on Agro-geoinformatics (Agro-geoinformatics)* (Hangzhou), 1–4.

James, J., and Maheshwar P, M. (2016). "Plant growth monitoring system, with dynamic user-interface," in *2016 IEEE Region 10 Humanitarian Technology Conference (R10-HTC)* (Agra), 1–5.

Li, Y., Dai, W., Zou, J., Xiong, H., and Zheng, Y. F. (2020). Scalable structured compressive video sampling with hierarchical subspace learning. *IEEE Trans. Circuits Syst. Video Technol.* 30, 3528–3543. doi: 10.1109/TCSVT.2019.2939370

Liu, Y., Yuan, X., Suo, J., Brady, D. J., and Dai, Q. (2019). Rank minimization for snapshot compressive imaging. *IEEE Trans. Pattern Anal. Mach. Intell.* 41, 2990–3006. doi: 10.1109/TPAMI.2018.2873587

Mun, S., and Fowler, J. E. (2012). "Dpcm for quantized block-based compressed sensing of images," in *2012 Proceedings of the 20th European Signal Processing Conference (EUSIPCO)* (Bucharest), 1424–1428.

Okayasu, T., Nugroho, A. P., Sakai, A., Arita, D., Yoshinaga, T., Taniguchi, R.-I., et al. (2017). "Affordable field environmental monitoring and plant growth measurement system for smart agriculture," in *2017 Eleventh International Conference on Sensing Technology (ICST)* (Sydney, NSW), 1–4.

Palangi, H., Ward, R., and Deng, L. (2016). Distributed compressive sensing: a deep learning approach. *IEEE Trans. Signal Process.* 64, 4504–4518. doi: 10.1109/TSP.2016.2557301

Peng, Y., Yang, M., Zhao, G., and Cao, G. (2022). Binocular-vision-based structure from motion for 3-d reconstruction of plants. *IEEE Geosci. Remote Sens. Lett.* 19, 1–5. doi: 10.1109/LGRS.2021.3105106

Piermattei, L., Karel, W., Wang, D., Wieser, M., Mokroš, M., Surový, P., et al. (2019). Terrestrial structure from motion photogrammetry for deriving forest inventory data. *Remote Sens.* 11, 950. doi: 10.3390/rs11080950

Prades-Nebot, J., Ma, Y., and Huang, T. (2009). "Distributed video coding using compressive sampling," in *2009 Picture Coding Symposium* (Chicago, IL), 1–4.

Qiu, W., Zhou, J., Zhao, H., and Fu, Q. (2015). Three-dimensional sparse turntable microwave imaging based on compressive sensing. *IEEE Geosci. Remote Sens. Lett.* 12, 826–830. doi: 10.1109/LGRS.2014.2363238

Rayhana, R., Xiao, G. G., and Liu, Z. (2021). Printed sensor technologies for monitoring applications in smart farming: a review. *IEEE Trans. Instrum. Meas.* 70, 1–19. doi: 10.1109/TIM.2021.3112234

Romano, Y., and Elad, M. (2016). Con-patch: when a patch meets its context. *IEEE Trans. Image Process.* 25, 3967–3978. doi: 10.1109/TIP.2016.2576402

Sajith V. V. V., Gopalakrishnan E. A., Sowmya V., and Soman K. P. (2019). "A complex network approach for plant growth analysis using images," in *2019 International Conference on Communication and Signal Processing (ICCSP)* (Chennai), 0249–0253.

Shechtman, E., and Irani, M. (2007). "Matching local self-similarities across images and videos," in *2007 IEEE Conference on Computer Vision and Pattern Recognition* (Minneapolis, MN), 1–8.

Somov, A., Shadrin, D., Fastovets, I., Nikitin, A., Matveev, S., Seledets, I., et al. (2018). Pervasive agriculture: iot-enabled greenhouse for plant growth control. *IEEE Pervasive Comput.* 17, 65–75. doi: 10.1109/MPRV.2018.2873849

Sullivan, G. J., Ohm, J.-R., Han, W.-J., and Wiegand, T. (2012). Overview of the high efficiency video coding (hevc) standard. *IEEE Trans. Circuits syst. Video Technol.* 22, 1649–1668. doi: 10.1109/TCSVT.2012.2221191

Tachella, J., Altmann, Y., Márquez, M., Arguello-Fuentes, H., Tourneret, J.-Y., and McLaughlin, S. (2020). Bayesian 3d reconstruction of subsampled multispectral single-photon lidar signals. *IEEE Trans. Comput. Imag.* 6, 208–220. doi: 10.1109/TCI.2019.2945204

Taimori, A., and Marvasti, F. (2018). Adaptive sparse image sampling and recovery. *IEEE Trans. Comput. Imag.* 4, 311–325. doi: 10.1109/TCI.2018.2833625

Tramel, E. W., and Fowler, J. E. (2011). "Video compressed sensing with multihypothesis," in *2011 Data Compression Conference* (Snowbird, UT), 193–202.

Tran, D. T., Yamaç, M., Degerli, A., Gabbouj, M., and Iosifidis, A. (2021). Multilinear compressive learning. *IEEE Trans. Neural Netw. Learn. Syst.* 32, 1512–1524. doi: 10.1109/TNNLS.2020.2984831

Trevisi, M., Akbari, A., Trocan, M., Rodríguez-Vázquez, A., and Carmona-Galán, R. (2020). Compressive imaging using rip-compliant cmos imager architecture and landweber reconstruction. *IEEE Trans. Circuits Syst. Video Technol.* 30, 387–399. doi: 10.1109/TCSVT.2019.2892178

Unde, A. S., and Pattathil, D. P. (2020). Adaptive compressive video coding for embedded camera sensors: compressed domain motion and measurements estimation. *IEEE Trans. Mob. Comput.* 19, 2250–2263. doi: 10.1109/TMC.2019.2926271

Yang, Y., Sun, J., Li, H., and Xu, Z. (2020). Admm-csnet: a deep learning approach for image compressive sensing. *IEEE Trans. Pattern Anal. Mach. Intell.* 42, 521–538. doi: 10.1109/TPAMI.2018.2883941

Yu, Y., Wang, B., and Zhang, L. (2010). Saliency-based compressive sampling for image signals. *IEEE Signal Process. Lett.* 17, 973–976. doi: 10.1109/LSP.2010.2080673

Zammit, J., and Wassell, I. J. (2020). Adaptive block compressive sensing: Toward a real-time and low-complexity implementation. *IEEE Access* 8, 120999–121013. doi: 10.1109/ACCESS.2020.3006861

Zhang, J., Zhao, D., and Jiang, F. (2013). "Spatially directional predictive coding for block-based compressive sensing of natural images," in *2013 IEEE International Conference on Image Processing* (Melbourne, VIC), 1021–1025.

Zhang, M., Wang, X., Chen, X., and Zhang, A. (2018). The kernel conjugate gradient algorithms. *IEEE Trans. Signal Process.* 66, 4377–4387. doi: 10.1109/TSP.2018.2853109

Zhang, P., Gan, L., Sun, S., and Ling, C. (2015). Modulated unit-norm tight frames for compressed sensing. *IEEE Trans. Signal Process.* 63, 3974–3985. doi: 10.1109/TSP.2015.2425809

Zhang, R., Wu, S., Wang, Y., and Jiao, J. (2020). High-performance distributed compressive video sensing: Jointly exploiting the hevc motion estimation and the ℓ_1-ℓ_1 reconstruction. *IEEE Access* 8, 31306–31316. doi: 10.1109/ACCESS.2020.2973392

Zhao, Z., Xie, X., Liu, W., and Pan, Q. (2020). A hybrid-3d convolutional network for video compressive sensing. *IEEE Access* 8, 20503–20513. doi: 10.1109/ACCESS.2020.2969290

Zhen, C., De-rong, C., and Jiu-lu, G. (2020). "A deep learning based distributed compressive video sensing reconstruction algorithm for small reconnaissance uav," in *2020 3rd International Conference on Unmanned Systems (ICUS)* (Harbin), 668–672.

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: 14 April 2022
doi: 10.3389/fpls.2022.877120

Ecology and Biodiversity Ontology Alignment for Smart Environment *via* Adaptive Compact Evolutionary Algorithm

Xingsi Xue[1,2]* *and Pei-Wei Tsai*[3]

[1] *Fujian Provincial Key Laboratory of Big Data Mining and Applications, Fujian University of Technology, Fuzhou, China,* [2] *Intelligent Information Processing Research Center, Fujian University of Technology, Fuzhou, China,* [3] *Department of Computer Science and Software Engineering, Swinburne University of Technology, Hawthorn, VIC, Australia*

Smart Environment (SE) focuses on the initiatives for healthy living, where ecological issues and biodiversity play a vital role in the environment and sustainability. To manage the knowledge on ecology and biodiversity and preserve the ecosystem and biodiversity simultaneously, it is necessary to align the data entities in different ecology and biodiversity ontologies. Since the problem of Ecology and Biodiversity Ontology Alignment (EBOA) is a large-scale optimization problem with sparse solutions, finding high-quality EBOA is an open challenge. Evolutionary Algorithm (EA) is a state-of-the-art technique in the ontology aligning domain, and this study further proposes an Adaptive Compact EA (ACEA) to address the problem of EBOA, which uses semantic reasoning to reduce searching space and adaptively guides searching direction to improve the algorithm's performance. In addition, we formally model the problem of EBOA as a discrete optimization problem, which maximizes the alignment's completeness and correctness through determining an optimal entity corresponding set. After that, a hybrid entity similarity measure is presented to distinguish the heterogeneous data entities, and an ACEA-based aligning technique is proposed. The experiment uses the famous Biodiversity and Ecology track to test ACEA's performance, and the experimental results show that ACEA-based aligning technique statistically outperforms other EA-based and state-of-the-art aligning techniques.

Keywords: ecology ontology, biodiversity ontology, ontology alignment, adaptive compact evolutionary algorithm, semantic reasoning

OPEN ACCESS

Edited by:
Yu Xue,
Nanjing University of Information Science and Technology, China

Reviewed by:
Ying Huo,
Nanjing Institute of Technology (NJIT), China
Xuejian Zhao,
Nanjing University of Posts and Telecommunications, China
Li Chaorong,
Yibin University, China

***Correspondence:**
Xingsi Xue
jack8375@gmail.com

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

Received: *16 February 2022*
Accepted: *14 March 2022*
Published: *14 April 2022*

Citation:
Xue X and Tsai P-W (2022) Ecology and Biodiversity Ontology Alignment for Smart Environment via Adaptive Compact Evolutionary Algorithm. Front. Plant Sci. 13:877120. doi: 10.3389/fpls.2022.877120

1. INTRODUCTION

Smart Environment (SE) is a knowledge-based system that focuses on initiatives for healthy living with an emphasis on environment and sustainability, where ecological issues and biodiversity play a vital role in urban citizens' life (Kumar, 2020). In particular, ecology research focuses on the ecosystems, habitat restoration practices, and communities of interest, which is of help to further our understanding of the environment and opportunities to affect change; biodiversity investigates the species' variability as well as their relationship to the environment. Since ecology and biodiversity are the most complex entities on this planet, the corresponding knowledge is usually modeled with the ontology (Madin et al., 2008), which is a powerful domain knowledge

modeling technique (Berners-Lee et al., 2001). Currently, more and more ecology and biodiversity ontologies have been developed, such as Environment Ontology (ENVO) and Plant Trait Ontology (PTO). However, since they are developed and maintained independently, a concept might be defined with different contexts, granularities, and terminologies, yielding the ontology heterogeneity problem (Karam et al., 2020). Examples of heterogeneous ecology and biodiversity ontologies are shown in **Table 1**.

Therefore, to preserve the ecosystem and biodiversity simultaneously and manage the knowledge on ecology and biodiversity, it is necessary to link the data entities in different ecology and biodiversity ontologies, which is the so-called Ecology and Biodiversity Ontology Alignment (EBOA).

Aligning ecology and biodiversity ontologies aims at finding a 0–1 Aligning Matrix (AM), whose element denotes whether two corresponding entities (the source ontology's entities in row and the target ontology's entities in column) are mapped by 1 or not by 0. Since the scale of the ecology and biodiversity ontologies are usually large, and the constraint of single cardinality on the aligning result, the problem of EBOA needs to find a large-scale AM (the number of its row and column is large) with sparse solutions (most of its element values are 0). Due to the large search space and richness of semantic meaning on different data entities, it is a complex task of aligning ecology and biodiversity ontologies. In recent years, Evolutionary Algorithm (EA) (Mirjalili, 2019) has become a popular technique in the ontology aligning domain (Acampora et al., 2013; Xue et al., 2018). Due to the population-based evolving paradigm, the classic EA's searching performance is low in terms of memory consumption and run time. To improve the efficiency, a new category of EA with the name Compact EA (CEA) is presented, which uses compact encoding mechanism to describe the whole population with probability estimation. CEA mimics EA's searching process by simplifying the evolving operators, but it is easy to get stuck in the local optima especially when two ontologies' scale is large. To overcome this drawback, this study further proposes an Adaptive CEA (ACEA), which uses the semantic reasoning to filter the negative correspondences, and adaptively alters the algorithm's searching direction to explore the unknown region. In the following, we list the contributions of this study:

TABLE 1 | The examples of heterogeneous ecology and biodiversity ontologies.

ENVO ontology[a]	**SWEET ontology**[b]
Divergent tectonic movement	Plate divergence
Tectonic movement	Continental drift
FLOTO ontology[c]	**PTO ontology**[d]
Inflorescence absent	Inflorescenceless
Leaf alternate placement	Phyllotaxy

[a] *http://agroportal.lirmm.fr/ontologies/ENVO.*
[b] *https://bioportal.bioontology.org/ontologies/SWEET.*
[c] *http://agroportal.lirmm.fr/ontologies/FLOPO.*
[d] *http://agroportal.lirmm.fr/ontologies/TO.*

- The optimization model of the problem of EBOA is presented;
- A hybrid entity similarity measure is proposed to distinguish the heterogeneous ecology and biodiversity data entities;
- An ACEA-based aligning technique is proposed, which uses semantic reasoning to reduce searching space, and adaptively guides the searching direction to efficiently align the ecology and biodiversity ontologies.

The introduction process of this study is as follows: before defining the problem of EBOA and entity similarity measure (Section 3), the EA-based aligning techniques are overviewed (Section 2); after that, the problem-specific ACEA is presented (Section 4), followed by the experimental results (Section 5); and finally, we draw the conclusion on this article's study (Section 6).

2. EVOLUTIONARY ONTOLOGY ALIGNING TECHNIQUE

With the rapid development of ontology engineering, the scale of an ontology has grown from hundreds of entities to tens of thousands of entities, and the semantic representation of the entities also become more and more complex, which makes the determination of a high-quality ontology alignment become an open challenge (Shvaiko and Euzenat, 2011). Essentially, the ontology aligning problem can be regarded as an optimizing issue that aims at maximizing the quality of final alignment, and EA-based aligning techniques have become a popular methodology to address this problem.

The first EA-based ontology aligning technique is proposed by Martinez-Gil et al. (2008) which tries to find an optimal way of combining different similarity measures for determining the final alignment. Later on, researchers have done a lot to improve this category of EA-based aligning techniques. Based on this study, Ginsca and Iftene (2010) further optimize the threshold for filtering final alignment. Acampora et al. (2012) propose a Hybrid EA (HEA) to improve the efficiency of classic EA's performance. Alves et al. (2012) further use the instance-level information in an ontology to construct the similarity measure and then use HEA to combine it with others. Currently, it is necessary to enhance the performance of population-based EA in terms of running time and memory so as to address the large-scale aligning task, such as addressing the problem of EBOA where the ontology contains tens of thousands of entities. To this end, an efficiency improvement strategy should be introduced to improve classic EA's performance. The most popular way is the utilization of a compact encoding based evolving paradigm, which describes the population with a probability distribution, and on this basis, it approximates the classic EA's evolving process. The first CEA-based aligning technique is proposed in Xue et al. (2015), which executes the evolving process by one Probability Vector (PV). According to the experimental results, CEA is able to significantly reduce EA's running time and memory consumption without sacrificing the alignment's quality. Later on, a Hybrid CEA (HCEA) (Xue and Wang, 2015a) and a CEA with a Re-sample Inheritance Mechanism (RIM) (Xue and Liu, 2022) are respectively proposed to further enhance CEA's performance. To address the large-scale aligning task, a divide-and-conquer

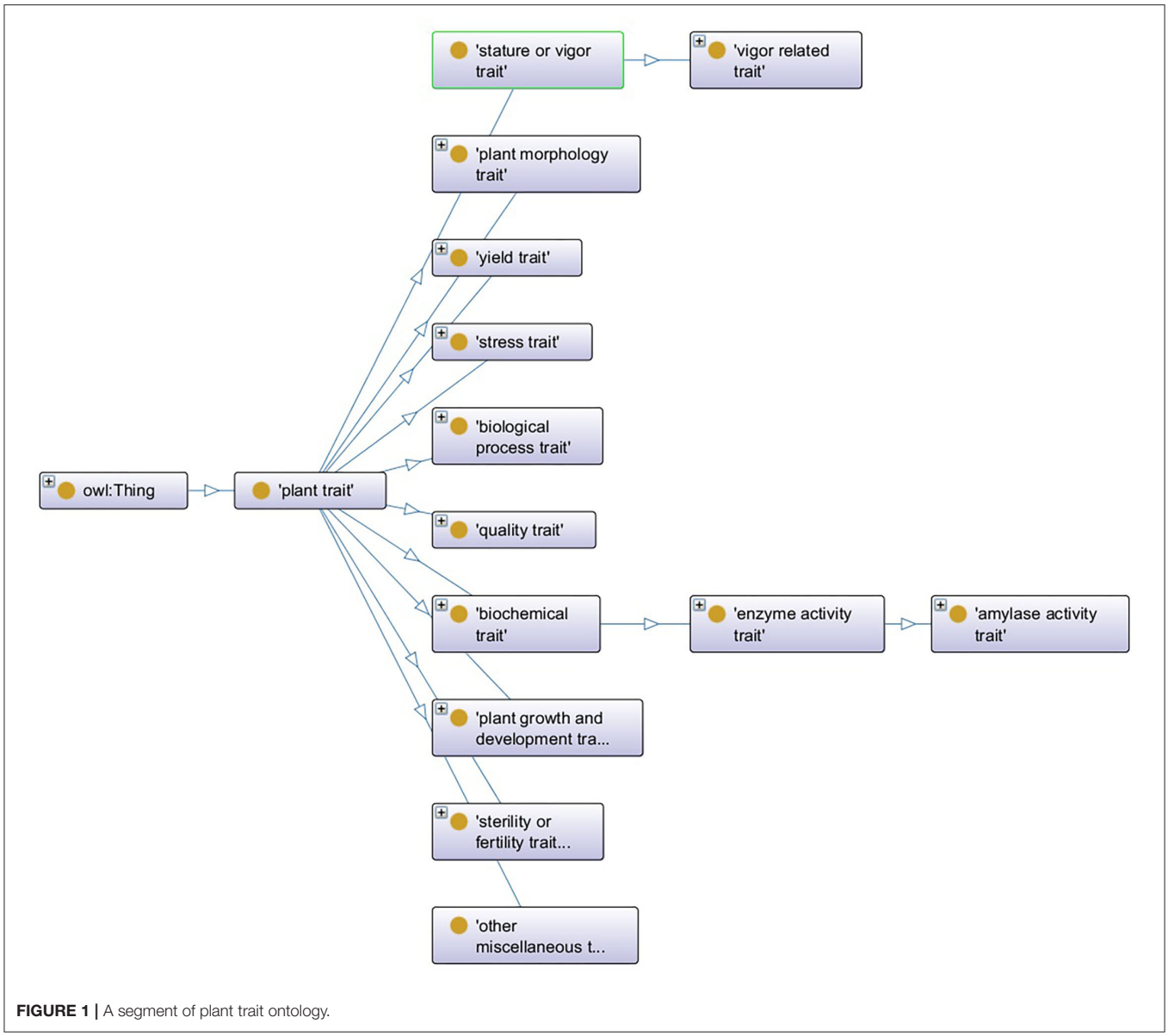


FIGURE 1 | A segment of plant trait ontology.

method is also presented, which is of help to reduce HCEA's searching space (Xue and Wang, 2015b; Xue and Zhang, 2021).

Existing EA-based aligning approaches need to maintain each similarity measure's corresponding AM, and on this basis, the optimization on the alignment can be executed, which greatly raises the computational complexity. In this study, we try to directly find a set of correspondences with the given similarity measure, which only needs to save several entity pairs' similarity value instead of maintaining all the similarity measures' corresponding entity pairs' similarity values. In addition, classic CEA only uses one PV to execute the optimizing process, which makes it easy to get stuck in the local optima when facing a complex optimization problem. To overcome this drawback, our approach proposes to adaptively maintain several Probability Matrices (PMs) to guide the algorithm's searching direction. Finally, since the problem of EBOA is a large-scale issue with sparse solutions, we propose semantic reasoning based initialization to reduce the algorithm's searching space and evenly distribute the AM's element values.

3. ECOLOGY AND BIODIVERSITY ONTOLOGY ALIGNMENT

3.1. The Problem of Ecology and Biodiversity Ontology Alignment

An ontology consists of the concepts, the datatype properties, and the object properties, which are referred to as entities (Xue et al., 2021). **Figure 1** shows a segment of PTO, where the oval symbol describes the concept's name, e.g., "plant trait," the arrow line is the object property or relationship between two concepts,

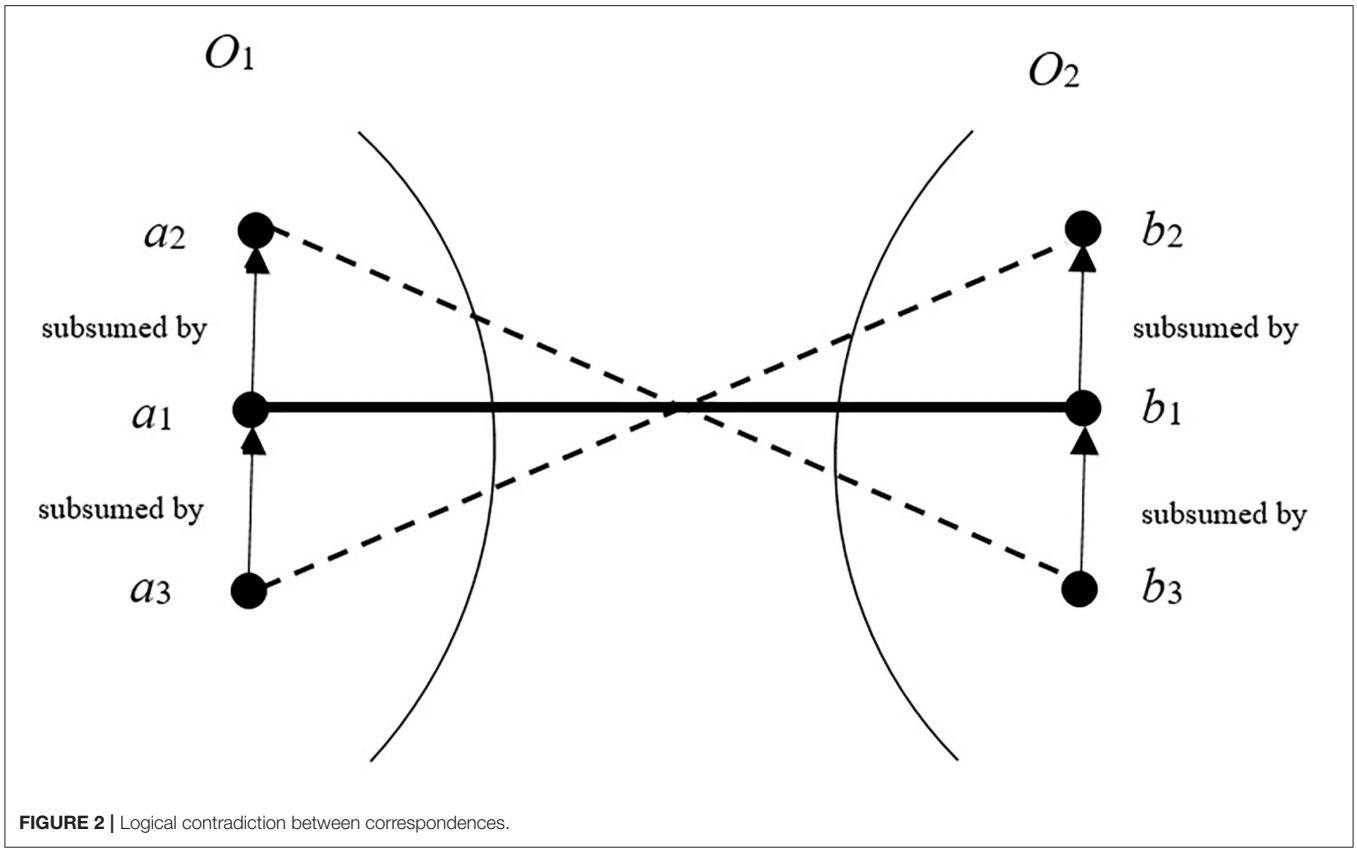


FIGURE 2 | Logical contradiction between correspondences.

e.g., the concept "quantity trait" is subsumed by "plant trait," and each concept has several datatype properties to describe its feature, e.g., the concept "plant trait" has the datatype property "definition" whose value is "A plant trait (TO:0000387) that is the commercial and /or economical value of the plant product, or its overall improvement."

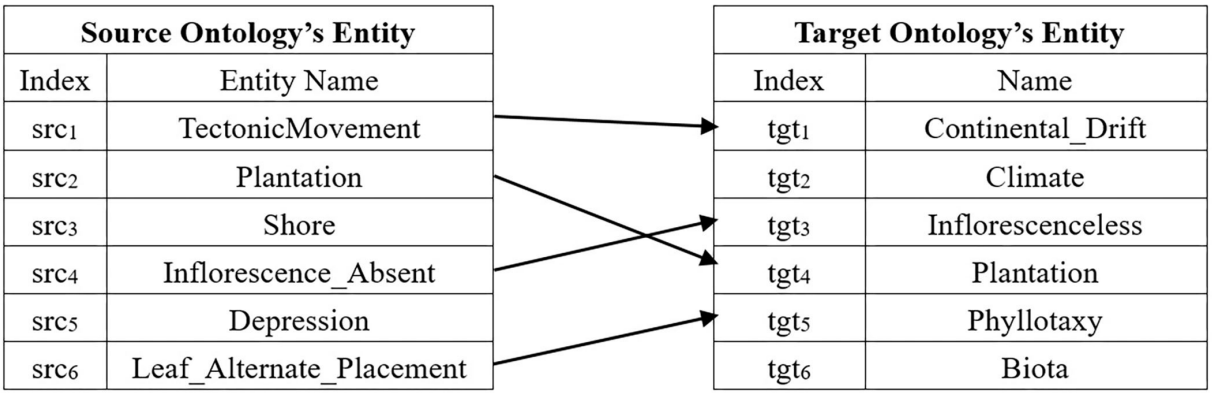
An entity correspondence consists of 4 elements, i.e., $e_1, e_2, rel, conf$, where e_1 and e_2 are respectively two ontologies' entities, *rel* is the type of their relationship (typically the equivalence $\equiv$), and *conf* denotes the confidence level that the correspondence holds, which is often measured by e_1 and e_2's similarity value. For example, in **Table 1**, $(ENVO : Tectonicmovement, SWEET : Continentaldrift, \equiv, 0.9)$ denote a correspondence between the concept "Tectonic movement" from ENVO and the concept "Continental drift" from SWEET, their relationship is the equivalence $\equiv$, and this correspondence's confidence value is 0.9. The ontology alignment is a correspondence set, whose quality is typically evaluated with recall, precision, and f-measure (Rijsberge, 1975). Since reference alignment is often not available in the practical aligning tasks, this study proposes three new metrics to approximate them. According to Wang et al. (2006), $recall(A)$ can be estimated by the number of correspondences found in A, i.e., $recall'(A) = norm(|A|)$, $precision(A)$ can be evaluated by the average similarity values of all the correspondences in A, i.e., $recall'(A) = \frac{\sum sim(corr_i)}{|A|}$ where $corr_i$ is i-th correspondence in A. To evaluate an alignment's quality in terms of both completeness and correctness, a comprehensive metric $f - measure'(A) = \frac{2 \times recall'(A) \times precision'(A)}{recall'(A) + precision'(A)}$ is presented, which calculate an alignment's harmony mean of its $recall'(A)$ and $precision'(A)$. Given two ontologies O_1 and O_2, a 0–1 matrix M and its corresponding alignment A, the problem of EBOA is defined as follows:

$$\begin{cases} max & f(M) \\ s.t. & M = [m_{i,j}]_{|O_1| \times |O_2|} \\ & m_{i,j} \in \{0,1\}, i = 1, 2, \cdots, |O_1| \end{cases} \quad (1)$$

where $|O_1|$ and O_2 are respectively O_1 and O_2's entity numbers, and $f(M)$ is equal to $f - measure'(A)$, and the model of EBOA aims at finding an optimal matrix by maximizing its corresponding alignment's $f - measure'$. In particular, the decision variable is a 0–1 matrix whose row and column are respectively two ontologies' entities, and its element value 1 means two corresponding entities are mapped, and 0 means not.

3.2. Entity Similarity Measure

The entity similarity measure calculates two entities' similarity value *conf*, which is a real number in [0,1]. $conf = 1$ means two entities are the same, and $conf = 0$ means they are totally different. To improve the result's confidence, usually, it is necessary to comprehensively consider three categories



Source Ontology's Entity	
Index	Entity Name
src_1	TectonicMovement
src_2	Plantation
src_3	Shore
src_4	Inflorescence_Absent
src_5	Depression
src_6	Leaf_Alternate_Placement

Target Ontology's Entity	
Index	Name
tgt_1	Continental_Drift
tgt_2	Climate
tgt_3	Inflorescenceless
tgt_4	Plantation
tgt_5	Phyllotaxy
tgt_6	Biota

The Corresponding Aligning Matrix

	tgt_1	tgt_2	tgt_3	tgt_4	tgt_5	tgt_6
src_1	1	0	0	0	0	0
src_2	0	0	0	1	0	0
src_3	0	0	0	0	0	0
src_4	0	0	1	0	0	0
src_5	0	0	0	0	0	0
src_6	0	0	0	0	1	0

FIGURE 3 | An example of encoding mechanism.

of similarity measures, i.e., string-based, linguistic-based, and structure-based similarity measures (Xue and Huang, 2022). To this end, this study proposes a hybrid entity similarity measure to comprehensively calculate the similarity value: (1) given two entities e_1 and e_2, before calculating their similarity value, the numbers, punctuations, and stop-words in their names are first removed; (2) the strings are split into the words, which are further lemmatized and stemmed; (3) in each word set, the word will be removed if it is the same literally or synonymous to the other one in Wordnet (Miller, 1995), and we obtain two word sets s_1 and s_2; finally, e_1 and e_2's similarity value is the same as the similarity value of two string s_1 and s_2:

$$sim(s_1, s_2) = \frac{\sum_{i=1}^{|W_1|} \max_{j=1\cdots|W_2|}(sim(w_{1,i}, w_{2,j})) + \sum_{j=1}^{|W_2|} \max_{i=1\cdots|W_1|}(sim(w_{1,i}, w_{2,j}))}{|W_1| + |W_2|} \tag{2}$$

where $|W_1|$ and $|W_1|$ are respectively the numbers of words in W_1 and W_2, and $w_{1,i}$ and $w_{2,j}$ are respectively the ith and jth words in W_1 and W_2; and $sim(w_{1,i}, w_{2,j})$ is calculated with Wordnet and N-gram measure (Kondrak, 2005):

$$sim(w_{1,i}, w_{2,j}) = \begin{cases} 1, & \text{two words are synonyms in Wordnet} \\ N\text{-}gram(w_{1,i}, w_{2,j}), & \text{otherwise} \end{cases} \tag{3}$$

4. ADAPTIVE COMPACT EVOLUTIONARY ALGORITHM

Adaptive compact evolutionary algorithm adaptively maintains PMs according to the current generation's population information, which is able to help the algorithm effectively exploit the unexplored domains. In addition, ACEA uses the anchor-based semantic reasoning strategy to initialize the individual and refine the new individuals, which can effectively reduce the algorithm's searching domain. The framework of ACEA is presented in **Algorithm 1**, which takes as input two ontologies to be aligned, and the output the alignment with best fitness value.

In the next, we successively present the Semantic Reasoning Based Initialization and adaptive PM maintenance.

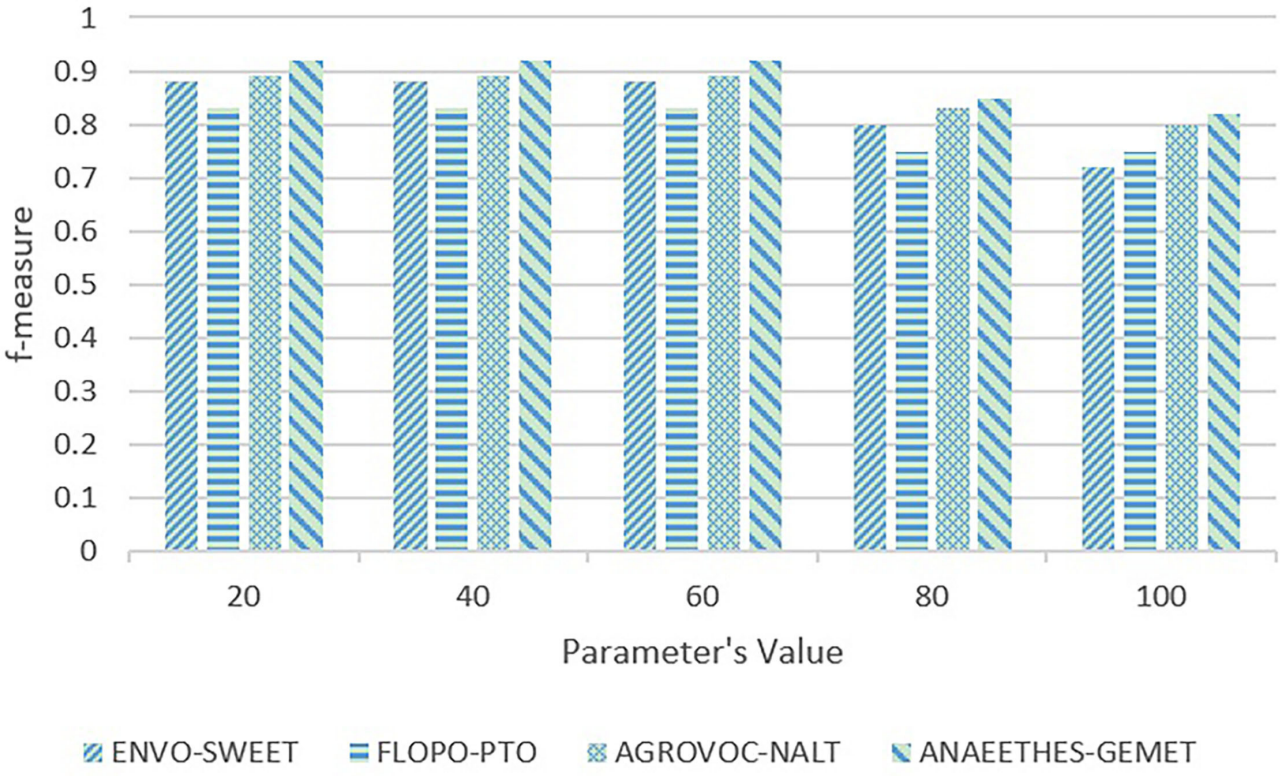


FIGURE 4 | Sensitivity testing on Adaptive Compact EA's (ACEA) parameter δ in terms of f-measure.

Algorithm 1 The Framework of Adaptive Evolutionary Algorithm

PM_{num} =2; //Initialize the number of PM
for $i = 0; i < PM_{num}; i$++ **do**
  $AM^i_{elite} = initializeAM(PM^i)$; //Initialize the elite AM
end for
$gen = 0$;
while $gen < maxGen$ **do**
  for $i = 0; i < PM_{num}; i$++ **do**
    $Update(PM^i)$; //Execute the evolutionary operator and update PM
  end for
  $PM_{num} = adaptivePMMaintenance()$; //Adaptively maintain the PMs
  $gen = gen + 1$;
end while

4.1. Semantic Reasoning Based Initialization

Typically, the correspondence with a high confidence value is referred to as Positive Anchor (PA), and the one with a low confidence value is called the Negative Anchor (NA) (Wang, 2010). The concepts in an ontology are modeled with the hierarchy graph (Chu et al., 2020), where the node denotes the concept and the edge represents the relationships between two concepts. **Figure 2** shows an example of correspondences' logical contradiction. As shown in the figure, the entities a_1, a_2, and a_3 are three entities of ontology O_1, and the entities b_1, b_2, and b_3 belong to ontology O_2. In O_1 (O_2), a_3 (b_3) is subsumed by a_1 (b_1), and a_1 (b_1) is subsumed by a_2 (b_2). Assuming the correspondence (a_1, b_1) is a PA, the correspondences (a_2, b_3) and (a_3, b_2) logically contradict with (a_1, b_1). It is obvious that the correspondences that contradict with some PA will not hold, and the confidence of correspondences that are in line with some NA should be reduced. According to this reasoning rule, the searching space of the algorithm can be reduced.

Each individual of ACEA is represented by a 0–1 matrix, i.e., the so-called AM. **Figure 3** shows an example of an encoding mechanism, wherein the top of the figure is a real alignment, and its corresponding AM is given below it. ACEA uses Probability Matrix (PM) to approximately describe the population, which has the same size as AM. PM's elements are the real number in [0,1], which denotes the probability of being 1 with respect to the corresponding gene bit. Therefore, we can use PM to generate AM by comparing its elements with a random number in [0,1].

Since the problem of EBOA is a large-scale issue with sparse solutions, it is necessary to evenly distribute the gene value

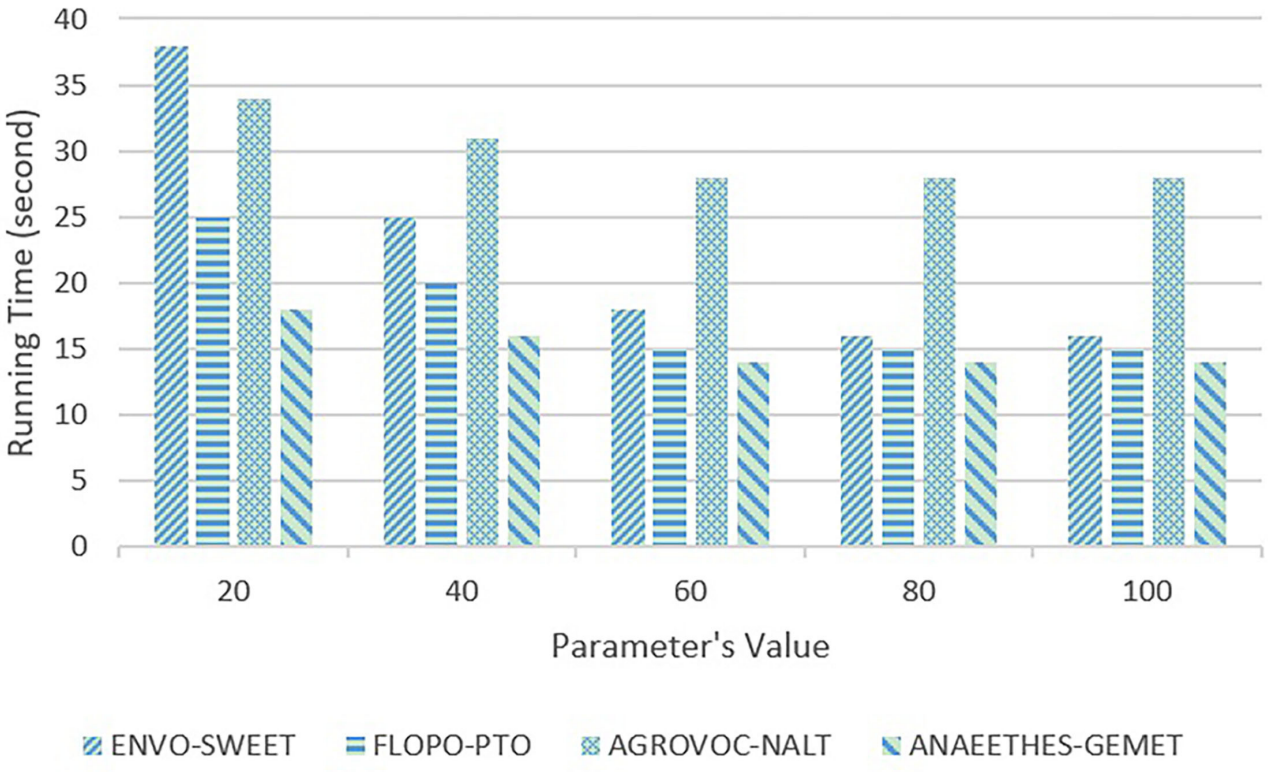


FIGURE 5 | Sensitivity testing on ACEA's parameter δ in terms of running time (second).

TABLE 2 | Comparisons among Adaptive Compact EA (ACEA), Compact EA (CEA), and Hybrid CEA (HCEA) in terms of mean f-measure and standard deviation.

Testing case	CEA *f-measure* (*stdDev*)	HCEA *f-measure* (*stdDev*)	ACEA *f-measure* (*stdDev*)
ENVO-SWEET	0.79 (0.03)	0.83 (0.01)	0.88 (0.01)
FLOPO-PTO	0.76 (0.01)	0.76 (0.02)	0.83 (0.01)
AGROVOC-NALT	0.78 (0.03)	0.80 (0.01)	0.89 (0.02)
ANAEETHES-GEMET	0.74 (0.02)	0.82 (0.01)	0.92 (0.01)

TABLE 3 | *T*-test on alignment's quality.

Testing case	(CEA, ACEA) *t*-value (*p*-value)	(HCEA, ACEA) *t*-value (*p*-value)
ENVO-SWEET	–5.69 (0.014)	–7.07 (0.009)
FLOPO-PTO	–9.89 (0.005)	–6.26 (0.012)
AGROVOC-NALT	–6.10 (0.012)	–8.04 (0.007)
ANAEETHES-GEMET	–16.09 (0.001)	–4.14 (0.026)

Algorithm 2 Initialization

```
Initialize Aligning Matrix AM by setting all its elements as 0;
if Probability Matrix PM is not given then
  initialize all the elements in PM as 0.5;
end if
initialize the Positive Anchor Set PAS;
for i = 0; i < PAS.length; i++ do
  for j = 0; j < AM.row; j++ do
    for k = 0; k < AM.column; k++ do
      if (entity_i, entity_j) is contradicted with PAS_i then
        PM_{i,j} = 0;
      else
        if random(0, 1) < PM_{i,j} then
          AM_{i,j} = 1;
        end if
      end if
    end for
  end for
end for
```

when initializing AM. **Algorithm 2** shows the pseudo-code of initialization.

We first initialize *AM* by setting all its elements as 0 and determine the positive anchor set *PAS* with the similarity measure. If probability matrix *PM* is not given, all its elements will be initialized as 0.5. Then, we compare all the correspondences in AM with *PAS*. If the correspondence is logically conflicted with *PAS*'s correspondence, its *AM* and *PM*'s values will be set as 0, otherwise, its value is decided by

comparing its corresponding *PM*'s value with a random number in [0,1]. Through semantic reasoning, the searching space can be significantly reduced, and initializing *PM*'s elements as 0.5 is also of help to ensure the even distribution of the gene values.

4.2. Updating Probability Matrix

Adaptive compact evolutionary algorithm combines the mechanisms of a classic EA with a competitive learning mechanism, which is effective to lead the algorithm to determine the optimal solution. To be specific, ACEA first uses its fitness function to evaluate its solution's fitness value by calculating its corresponding alignment's $f - measure'$, and through competitions between the individuals, the algorithm updates PM by moving it toward the elite individual. The process of updating PM is presented in **Algorithm 3**.

Algorithm 3 Updating Probability Matrix

$AM^{new} = PM.generateAM()$;
$AM' = PM.generateAM()$;
for $i = 0; i < AM^{new}.row; i$++ **do**
  for $j = 0; j < AM^{new}.column; j$++ **do**
    if $random(0,1) < 0.5$ **then**
      $AM^{new}_{i,j} = AM'_{i,j}$;
    end if
  end for
end for
$compete(AM^{new}, AM^{elite})$;
if $winner == AM^{new}$ **then**
  $AM^{elite} = AM^{new}$;
end if
for $i = 0; i < PM.row; i$++ **do**
  for $j = 0; j < PM.column; j$++ **do**
    if $AM^{elite}_{i,j} == 1$ **then**
      $PM^{elite}_{i,j} += 0.01$;
    else
      $PM^{elite}_{i,j} -= 0.01$;
    end if
  end for
end for
for $i = 0; i < PM^{elite}.row; i$++ **do**
  for $j = 0; j < PM^{elite}.column; j$++ **do**
    if $corr(e_i, e_j).conf < 0.2$ **then**
      $PM^{elite}_{i,j} = 0$;
      if $corr(e_m, e_n)$ is $corr(e_i, e_j)$'s neighbor correspondence **then**
        if $corr(e_m, e_n)$ does not logically contradict with $corr(e_i, e_j)$ **then**
          $PM^{elite}_{m,n} -= 0.01$;
        end if
      end if
    end if
  end for
end for

In **Algorithm 3**, we first generate two AMs and use them to obtain a new AM AM^{new} with the uniform crossover operator. Then, AM^{new} is compared with the elite AM AM^{elite}, and the winner will become the elite AM. After that, we use AM^{elite} to update its corresponding PM: if $AM^{elite}_{i,j} == 1$, then $PM^{elite}_{i,j} += 0.01$; otherwise, $PM^{elite}_{i,j} -= 0.01$. We update PM so that the newly generated AM will be closer to the elite AM. Finally, we find the NA $corr(e_i, e_j)$ from AM^{elite}, and their neighbor correspondence $corr(e_m, e_n)$ where the shortest path between e_m (or e_n) and e_i (or e_j) in the ontology hierarchy graph is less than 2, we pick up those do logically contradict with $corr(e_i, e_j)$ and reduce their corresponding *PM* elements' values by 0.01. In particular, the step length of updating *PM* determines the algorithm's learning rate. If the step length is too large, the algorithm converges quickly, i.e., the value of *PM*âĂŸs elements are close to 1 or 0; and if it is too small, the algorithm consumes a long running time. Here, we empirically set the step length as 0.01, which is able to ensure the highest average quality of alignments on all testing cases.

4.3. Adaptive Probability Matrix Maintenance

At the end of the generation, adaptive PM maintenance is executed to adjust the algorithm's

Algorithm 4 Adaptive Probability Matrix Maintenance

if $\sum(|PM^a_{i,j} - PM^b_{i,j}|) < 0.5$ **then**
  if PM^a_{elite} is better than PM^b_{elite} **then**
    remove PM^b;
  else
    remove PM^a;
  end if
end if
if All elite AMs keep unchanged for δ generations **then**
  for $i = 0; i < PM^{new}.row; i$++ **do**
    for $j = 0; j < PM^{new}.column; j$++ **do**
      if $PM^{max}_{i,j} = PM^{min}_{i,j} = 1 or 0$ **then**
        $PM^{new}_{i,j} = 1 or 0$;
      end if
      if $PM^{max}_{i,j} < 0.5$ **then**
        $PM^{new}_{i,j} = PM^{max}_{i,j} + rand(0,1)(1 - PM^{max}_{i,j})$;
      end if
      if $PM^{min}_{i,j} > 0.5$ **then**
        $PM^{new}_{i,j} = (1 - rand(0,1))PM^{min}_{i,j}$;
      end if
      if $PM^{max}_{i,j} > 0.5$ and $PM^{min}_{i,j} < 0.5$ **then**
        $PM^{new}_{i,j} = 0.5$;
      end if
    end for
  end for
  $AM^{new}_{elite} = initializeAM(PM^{new})$;
end if

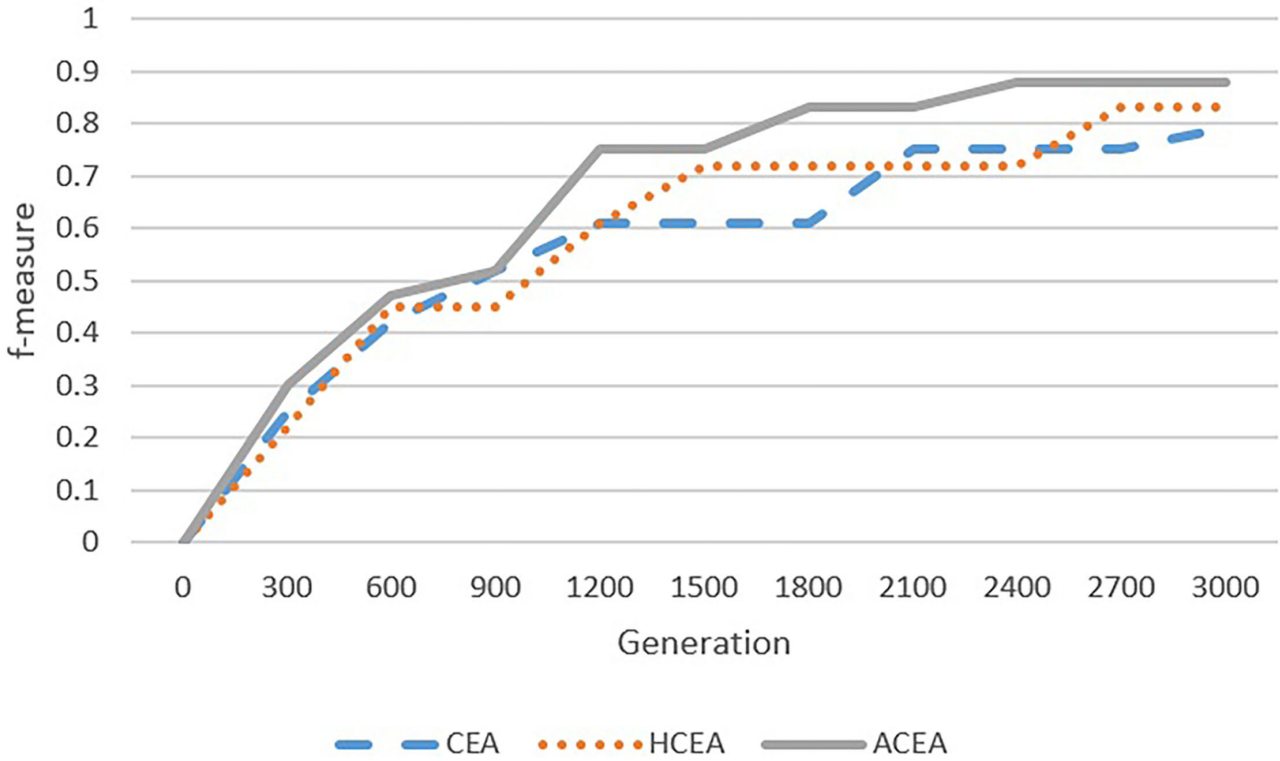


FIGURE 6 | Comparison among ACEA, CEA, and HCEA in terms of convergence graph on ENVO-SWEET.

searching direction. The pseudo-code of adaptive population maintenance is shown in **Algorithm 4**.

In **Algorithm 4**, we first calculate the distance between the existing PMs. The smaller distance indicates a larger overlap between their searching directions, and therefore, the one with worse elite AM should be deleted. When all elite AMs keep unchanged for δ generations, i.e., the algorithm gets stuck in the local optima, we add a new PM PM^{new} with a different searching direction. To determine PM^{new}'s elements, we need to analyze the existing PMs' elements distribution through the maximum probability and minimum probability. To be specific, for each of PM^{new}'s element $PM^{new}_{i,j}$, we find its corresponding maximum probability $PM^{max}_{i,j}$ and minimum probability $PM^{min}_{i,j}$ from the existing PMs. If they are all equal to 1 or 0, we will set $PM^{new}_{i,j}$ as 1 or 0. If $PM^{max}_{i,j} < 0.5$, $PM^{new}_{i,j}$ will be put in the left of all existing probabilities, i.e., $PM^{new}_{i,j} = PM^{max}_{i,j} + rand(0,1)(1 - PM^{max}_{i,j})$. If $PM^{min}_{i,j} > 0.5$, $PM^{new}_{i,j}$ will be put in the right of all existing probabilities, i.e., $PM^{new}_{i,j} = (1 - rand(0,1))PM^{min}_{i,j})$. If $PM^{max}_{i,j} > 0.5$ and $PM^{min}_{i,j} < 0.5$, $PM^{new}_{i,j}$ will be put in the middle, i.e., $PM^{new}_{i,j} = 0.5$. Finally, we initialize the elite AM AM^{new}_{elite} for PM^{new}.

5. EXPERIMENT

5.1. Experimental Configuration

We use the track of Biodiversity and Ecology in Ontology Alignment Evaluation Initiative (OAEI)[1] to test ACEA's performance. Biodiversity and Ecology track consists of four pairs of ontologies in the biodiversity and ecology domain: (1) ENVO[2]-SWEET[3], (2) PTO[4]-FLOPO[5], (3) AGROVOC[6]-NALT[7], (4) GEMET[8]-ANAEETHES[9]. All of these ontologies are widely used in various projects and researches on biodiversity and ecology, which are developed in parallel and are significantly overlapping.

In the experiment, we compare ACEA with CEA (Xue et al., 2015), HCEA (Xue and Chen, 2019), and OAEI's participants. In particular, CEA's configuration is as follows:

[1]http://oaei.ontologymatching.org
[2]http://agroportal.lirmm.fr/ontologies/ENVO
[3]https://bioportal.bioontology.org/ontologies/SWEET
[4]http://agroportal.lirmm.fr/ontologies/TO
[5]http://agroportal.lirmm.fr/ontologies/FLOPO
[6]http://agroportal.lirmm.fr/ontologies/AGROVOC
[7]http://agroportal.lirmm.fr/ontologies/NALT
[8]http://agroportal.lirmm.fr/ontologies/GEMET
[9]http://agroportal.lirmm.fr/ontologies/ANAEETHES

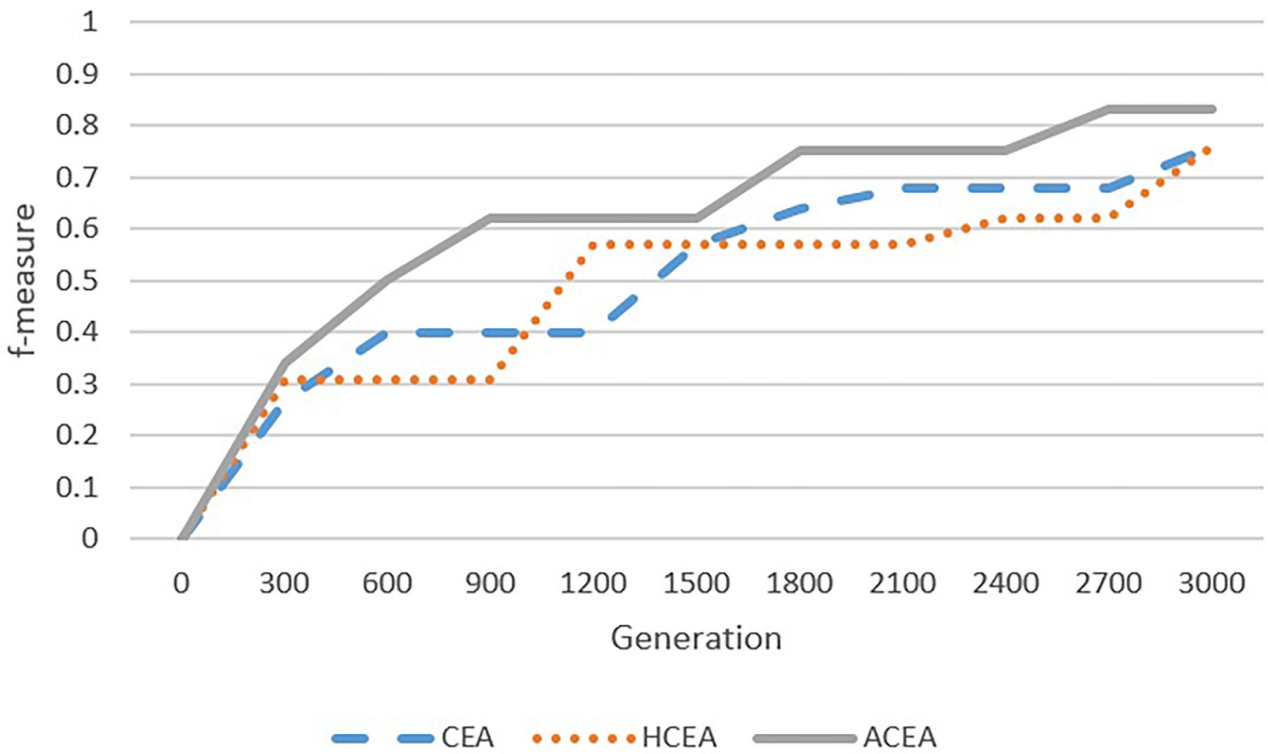


FIGURE 7 | Comparison among ACEA, CEA, and HCEA in terms of convergence graph on FLOPO-PTO.

- The maximum generation = 3,000;
- The step length for updating PV = 0.01.

The configuration of HCEA's is as following:

- The maximum generation = 3,000;
- The step length for updating PV = 0.01;
- The crossover probability = 0.6;
- The mutation probability = 0.03;
- The mutation shift = 0.05.

Additionally, ACEA's configuration is given as follows:

- The maximum generation = 3,000;
- The threshold for activate PM Maintenance = 60;
- The step length for updating probability matrix = 0.01.

Ontology Alignment Evaluation Initiative's participants' results are from OAEI's official website[10]. We first show the sensitivity testing on ACEA's parameter, then ACEA is compared with CEA and HCEA in terms of f-measure and convergence graph, and finally, ACEA is compared with the state-of-the-art ontology aligning techniques. ACEA, CEA, and HCEA's results are the mean value of 30 independent runs.

[10]http://oaei.ontologymatching.org/2021/biodiv/index.html

5.2. Experimental Results

First, the sensitivity testings are carried out on ACEA's parameter δ that determines the timing of executing the PM maintenance. If δ is too large, ACEA would get stuck in the local optima for a long time, which would hamper the algorithm from converging on the global optima, and if δ is too small, there would be too many PVs, which increases the computational complexity. We empirically take five representative values, i.e., 20, 40, 60, 80, and 100, to execute the sensitivity testing on δ, whose results are shown in **Figures 4, 5**.

In **Figures 4, 5**, with the increasing values, the quality of alignments begins to deteriorate when $\delta > 60$, and the running time start to decrease, and when $\delta = 60$, it reaches the bottom. Therefore, the parameter $\delta = 60$ is able to better trade-off the quality of alignments and the algorithm's running time.

In **Table 2**, we compare ACEA with CEA and HCEA with mean f-measure *f-measure* and the standard deviation *stdDev*. In **Table 2**, the statistical *t*-test (Schmetterer and Lehmann, 1962) is executed on the data presented in **Table 3**.

In **Table 3**, the T-test's degree of freedom of is 2, and the significance level is 0.05. On all testing cases, the *p*-values are all smaller than 0.05, and thus, we can draw the conclusion that ACEA statistically outperforms CEA and HCEA based aligning techniques at the significance level of 5%. It is obvious that through adaptively maintaining PMs, ACEA

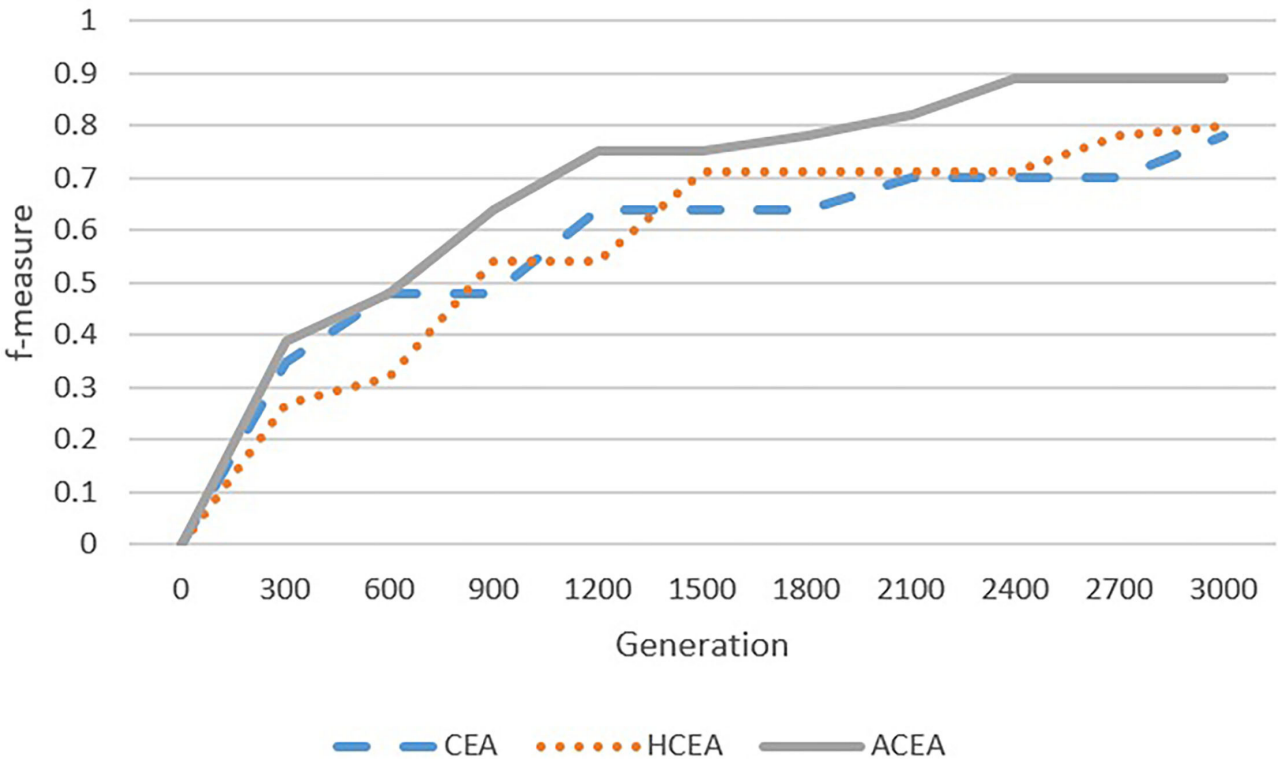


FIGURE 8 | Comparison among ACEA, CEA, and HCEA in terms of convergence graph on AGROVOC-NALT.

can effectively explore the whole feasible region and find high-quality alignments.

As depicted in **Figures 6–9**, with the introduction of semantic reasoning, the searching space of ACEA can be significantly reduced, which makes it able to more efficiently converge to the global optimal solution.

Finally, we compare ACEA with OAEI's participants on Biodiversity and Ecology track through f-measure. In **Table 4**, we can see that ACEA's mean f-measure outperforms all the state-of-the-art ontology aligning systems on all testing cases. ACEA makes use of the evolutionary paradigm to iteratively refine the alignment's quality, which is a more effective way of improving the alignment's quality than other machine learning based aligning approaches (such as ALOD2Vec, POMap, and DOME), logical reasoning based aligning methods (such as Lily, LogMap Family, and XMap) and Word-based aligning techniques (such as AML, Wiktionary, FCAMapKG, ATBox, and KGMatcher).

5.3. Discussions on Experimental Results

Compact evolutionary algorithm combines the mechanisms of a classic EA with competitive learning, which is effective to lead the algorithm to determine the optimal solution. In addition, the simplicity of CEA, which does not require all the mechanisms of an EA, rather the few steps in the algorithm are small and simple. HCEA further introduces local refinements on the elite solution, which allows increasing the convergence speed *via* the local search. Compared with CEA and HCEA, ACEA works based on the probabilistic modeling of promising solutions, which makes it easier to predict the movements of the populations in the search space. When confronted with complex optimization issues, ACEA is able to jump out of the local optima through adaptively PM maintenance, which guides the algorithm to explore the potential search space and learn a more complex probabilistic model. Therefore, ACEA outperforms CEA and HCEA in terms of both qualities of alignments and computational efficiency.

In addition, ACEA comprehensively aggregates three broad categories of entity similarity measure, i.e., syntactic-based similarity measure, linguistic-based similarity measure, and structure-based similarity measure, which lead to better alignments than the ones that only take into consideration one or two of them, such as AML, LogMap Family, POMap, XMap, DOME, FCAMapKG. This is because when facing a different heterogeneous situation, none of the similarity measures could be effective in all matching tasks, and taking into consideration more similarity measures could be of help to find the correct correspondences.

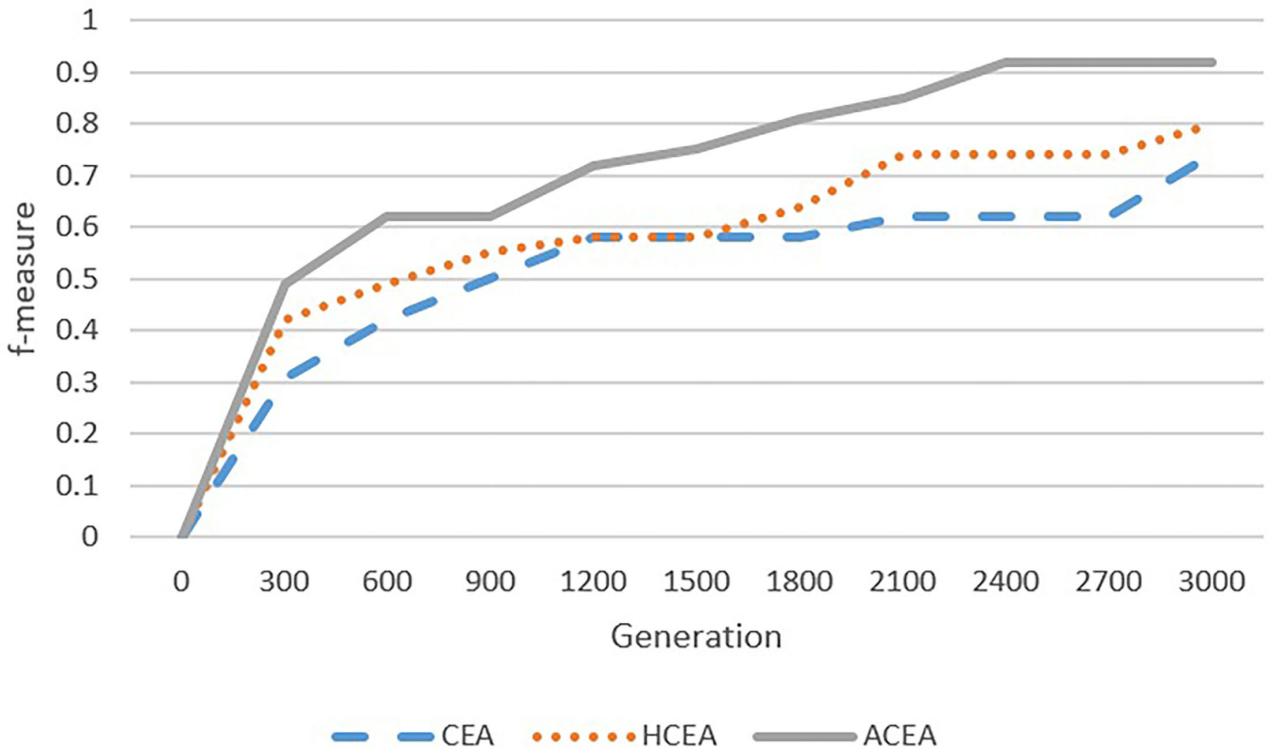


FIGURE 9 | Comparison among ACEA, CEA, and HCEA in terms of convergence graph on ANAEETHES-GEMET.

TABLE 4 | Comparisons among ACEA and state-of-the-art ontology aligning techniques in terms of f-measure.

Aligning system	ENVO-SWEET	FLOPO-PTO	AGROVOC-NALT	ANAEETHES-GEMET
AML (Lima et al., 2020)	0.84	0.86	0.87	0.85
Lily (Wu et al., 2019)	0.73	0.68	-	-
LogMap (Jiménez-Ruiz, 2020)	0.78	0.80	-	0.89
LogMapBio (Jiménez-Ruiz, 2020)	0.77	0.79	-	0.89
LogMapLite (Jiménez-Ruiz, 2020)	0.77	0.75	-	0.49
POMap (Laadhar et al., 2018)	0.78	0.68	-	-
XMap (Djeddi et al., 2015)	0.78	0.76	-	-
DOME (Hertling and Paulheim, 2018)	-	0.73	-	-
FCAMapKG (Algergawy et al., 2019)	0.63	0.69	-	-
POMap (Laadhar et al., 2018)	0.69	0.68	-	-
ATBox (Hertling and Paulheim, 2020)	0.69	0.71	-	-
Wiktionary (Portisch and Paulheim, 2020)	-	0.002	-	-
ALOD2Vec (Portisch et al., 2020)	-	0.002	-	0.10
KGMatcher (Fallatah et al., 2021)	0.005	-	-	0.063
ACEA	0.88	0.83	0.89	0.92

The symbol "-" means that the corresponding matching technique is not able to determine an alignment.

6. CONCLUSION

To manage knowledge on ecology and biodiversity and preserve the ecosystem and biodiversity simultaneously, it is necessary to link the data entities in different ecology and biodiversity ontologies. To this end, this study proposes an ACEA-based ecology and biodiversity ontology aligning technique. In particular, the problem of EBOA is modeled as a large-scale discrete optimization problem with a sparse solution. Then, a hybrid entity similarity measure is

presented to calculate the ecology and biodiversity entities' similarity. Finally, a problem-specific ACEA is proposed, which introduces semantic reasoning and adaptive PM maintenance to efficiently solve the problem of EBOA. The experimental results show that the evolutionary paradigm is able to find a better alignment than other artificial techniques and the proposed semantic reasoning and adaptive PM maintenance are able to further improve the algorithm's efficiency.

Although ACEA based aligning technique shows its superiority in the experiment, it is not able to detect the m:n correspondence, i.e., multiple source entities are mapped with multiple target entities, which is a common complex correspondence pattern. In addition, ACEA is also not able to find other semantic relationships among the entities, such as the subsumption. Finally, the divide-and-conquer approach has been proved to be a viable method that can facilitate the effectiveness of aligning process (Hu et al., 2008), and we are also interested in utilizing the ontology partitioning technique to pre-process two ontologies.

DATA AVAILABILITY STATEMENT

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

AUTHOR CONTRIBUTIONS

XX and P-WT proposed the idea. XX performed the literature review and comparative analyses and wrote the manuscript draft. P-WT performed the experiments and made the figures. All authors read the manuscript draft, commented on it, and confirmed it before submission.

FUNDING

This study was supported by the National Natural Science Foundation of China (No. 62172095), the Natural Science Foundation of Fujian Province (No. 2020J01875), and the Scientific Research Foundation of the Fujian University of Technology (No. GY-Z17162).

REFERENCES

Acampora, G., Loia, V., Salerno, S., and Vitiello, A. (2012). A hybrid evolutionary approach for solving the ontology alignment problem. *Int. J. Intell. Syst.* 27, 189–216. doi: 10.1002/int.20517

Acampora, G., Loia, V., and Vitiello, A. (2013). Enhancing ontology alignment through a memetic aggregation of similarity measures. *Inf. Sci.* 250, 1–20. doi: 10.1016/j.ins.2013.06.052

Algergawy, A., Faria, D., Ferrara, A., Fundulaki, I., Harrow, I., Hertling, S., et al. (2019). "Results of the ontology alignment evaluation initiative 2019," in *The Fourteenth International Workshop on Ontology Matching collocated with the 18th International Semantic Web Conference (ISWC 2019)*, *Vol.* 2536 (Auckland), 46–85.

Alves, A., Revoredo, K., and Bai ao, F. (2012). "Ontology alignment based on instances using hybrid genetic algorithm," in *Proceedings of the 7th International Conference on Ontology Matching-Vol. 946* (CEUR-WS. Org) (Boston), 242–243.

Berners-Lee, T., Hendler, J., Lassila, O. (2001). The semantic web. *Sci. Am.* 284, 28–37. doi: 10.1038/scientificamerican0501-34

Chu, S.-C., Xue, X., Pan, J.-S., and Wu, X. (2020). Optimizing ontology alignment in vector space. *J. Internet Technol.* 21, 15–22. doi: 10.3966/160792642020012101002

Djeddi, W. E., Khadir, M. T., and Yahia, S. B. (2015). "Xmap: results for oaei 2015," in *CEUR Workshop Proceedings* (Bethlehem), 216–221.

Fallatah, O., Zhang, Z., and Hopfgartner, F. (2021). "Kgmatcher results for oaei 2021," in *Proceedings of the 16th International Workshop on Ontology Matching co-located with the 20th International Semantic Web Conference (ISWC 2021)* (Virtual Conference), 160–166.

Ginsca, A.-L., and Iftene, A. (2010). "Using a genetic algorithm for optimizing the similarity aggregation step in the process of ontology alignment," in *9th Roedunet International Conference* (Sibiu), 118–122.

Hertling, S., and Paulheim, H. (2018). "Dome results for oaei 2018," in *The Thirteenth International Workshop on Ontology Matching collocated with the 17th International Semantic Web Conference (ISWC 2018)* (Monterey), *Vol.* 2288, 144–151.

Hertling, S., and Paulheim, H. (2020). "Atbox results for oaei 2020," in *The Fifteenth International Workshop on Ontology Matching Collocated With the 19th International Semantic Web Conference (ISWC 2020)*, *Vol. 2788* (Virtual Conference: RWTH), 168–175.

Hu, W., Qu, Y., and Cheng, G. (2008). Matching large ontologies: a divide-and-conquer approach. *Data Knowl. Eng.* 67, 140–160. doi: 10.1016/j.datak.2008.06.003

Jiménez-Ruiz, E. (2020). "Logmap family participation in the oaei 2020," in *The Fifteenth International Workshop on Ontology Matching collocated With the 19th International Semantic Web Conference (ISWC 2020)*, *Vol. 2788* (CEUR-WS), 201–203.

Karam, N., Khiat, A., Algergawy, A., Sattler, M., Weiland, C., and Schmidt, M. (2020). Matching biodiversity and ecology ontologies: challenges and evaluation results. *Knowl. Eng. Rev.* 35, 132. doi: 10.1017/S0269888920000132

Kondrak, G. (2005). "N-gram similarity and distance," in *International Symposium on String Processing and Information Retrieval* (Lille: Springer), 115–126.

Kumar, V. (2020). "Smart environment for smart cities," in *Smart Environment for Smart Cities* (Singapore: Springer), 1–53.

Laadhar, A., Ghozzi, F., Bousarsar, I. M., Ravat, F., Teste, O., and Gargouri, F. (2018). "Oaei 2018 results of pomap++," in *The Thirteenth International Workshop on Ontology Matching Collocated With the 17th International Semantic Web Conference (ISWC 2018)* (Monterey), 192–196.

Lima, B., Faria, D., Couto, F. M., Cruz, I. F., and Pesquita, C. (2020). "Oaei 2020 results for aml and amlc," in *The Fifteenth International Workshop on Ontology Matching collocated With the 19th International Semantic Web Conference (ISWC 2020)* (Virtual Conference), 154–160.

Madin, J.-S., Bowers, S., Schildhauer, M.-P., and Jones, M. B. (2008). Advancing ecological research with ontologies. *Trends Ecol. Evol.* 23, 159–168. doi: 10.1016/j.tree.2007.11.007

Martinez-Gil, J., Alba, E., and Aldana-Montes, J. F. (2008). "Optimizing ontology alignments by using genetic algorithms," in *Proceedings of the Workshop on Nature Based Reasoning for the Semantic Web* (Karlsruhe), 1–15.

Miller, G. A. (1995). Wordnet: A lexical database for english. *Commun. ACM* 38, 39–41. doi: 10.1145/219717.219748

Mirjalili, S. (2019). "Genetic algorithm," in *Evolutionary Algorithms and Neural Networks* (Cham: Springer), 43–55.

Portisch, J., Hladik, M., and Paulheim, H. (2020). "Alod2vec matcher results for oaei 2020," in *The Fifteenth International Workshop on Ontology Matching collocated With the 19th International Semantic Web Conference (ISWC 2020)*, *Vol. 2788* (Virtual Conference: RWTH), 147–153.

Portisch, J., and Paulheim, H. (2020). "Wiktionary matcher results for oaei 2020," in *The Fifteenth International Workshop on Ontology Matching collocated with the*

19th International Semantic Web Conference (ISWC 2020), *Vol. 2788* (Virtual Conference: RWTH), 225–232.

Rijsberge, C. J. V. (1975). *Information Retrieval*. London: University of Glasgow, Butterworth.

Schmetterer, L., and Lehmann, E. L. (1962). Testing statistical hypotheses. *Econometrica* 30, 462–465. doi: 10.2307/1909917

Shvaiko, P., and Euzenat, J. (2011). Ontology matching: state of the art and future challenges. *IEEE Trans. Knowl. Data Eng.* 25, 158–176. doi: 10.1109/TKDE.2011.253

Wang, J., Ding, Z., and Jiang, C. (2006). "Gaom: genetic algorithm based ontology matching," in *Proceedings of IEEE Asia8CPacific Conference on Services Computing* (Guangzhou: IEEE), 61–620.

Wang, P. (2010). "Lily-lom: an efficient system for matching large ontologies with non-partitioned method," in *Proceedings of CEUR Workshop, Vol. 658* (Shanghai: Citeseer), 69–72.

Wu, J., Pan, Z., Zhang, C., and Wang, P. (2019). "Lily results for oaei 2019," in *The Fourteenth International Workshop on Ontology Matching Collocated With the 18th International Semantic Web Conference (ISWC 2019)* (Auckland), 153–159.

Xue, X., and Chen, J. (2019). Optimizing ontology alignment through hybrid population-based incremental learning algorithm. *Mem. Comput*. 11, 209–217. doi: 10.1007/s12293-018-0255-8

Xue, X., Chen, J., and Yao, X. (2018). Efficient user involvement in semiautomatic ontology matching. *IEEE Trans. Emerg. Top. Comput. Intell.* 5, 214–224. doi: 10.1109/TETCI.2018.2883109

Xue, X., and Huang, Q. (2022). Generative adversarial learning for optimizing ontology alignment. *Expert Syst*. doi: 10.1111/exsy.12936. [Epub ahead of print].

Xue, X., Liu, J., Tsai, P.-W., Zhan, X., and Ren, A. (2015). "Optimizing ontology alignment by using compact genetic algorithm," in *2015 11th International Conference on Computational Intelligence and Security (CIS)* (Shenzhen: IEEE), 231–234.

Xue, X., and Liu, W. (2022). Integrating heterogeneous ontologies in asian languages through compact genetic algorithm with annealing re-sample inheritance mechanism. *ACM Trans. Asian Low-Res. Lang. Inform. Process*. 1–21. doi: 10.1145/3519298

Xue, X., and Wang, Y. (2015a). Optimizing ontology alignments through a memetic algorithm using both matchfmeasure and unanimous improvement ratio. *Artif. Intell.* 223, 65–81. doi: 10.1016/j.artint.2015.03.001

Xue, X., and Wang, Y. (2015b). Using memetic algorithm for instance coreference resolution. *IEEE Trans. Knowl. Data Eng.* 28, 580–591. doi: 10.1109/TKDE.2015.2475755

Xue, X., Wu, X., Jiang, C., Mao, G., and Zhu, H. (2021). Integrating sensor ontologies with global and local alignment extractions. *Wireless Commun. Mobile Comput*. 2021, 1–10. doi: 10.1155/2021/6625184

Xue, X., and Zhang, J. (2021). Matching large-scale biomedical ontologies with central concept based partitioning algorithm and adaptive compact evolutionary algorithm. *Appl. Soft. Comput*. 106, 1–11. doi: 10.1016/j.asoc.2021.107343

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: 28 April 2022
doi: 10.3389/fpls.2022.818376

Evaluation of Ecological Suitability and Quality Suitability of *Panax notoginseng* Under Multi-Regionalization Modeling Theory

JiaQi Yue, ZhiMin Li, ZhiTian Zuo and YuanZhong Wang**

Medicinal Plants Research Institute, Yunnan Academy of Agricultural Sciences, Kunming, China

Panax notoginseng is an important medicinal plant in China, but there are some limitations in the ecological suitability study, such as incomplete investigation of species distribution, single regionalization modeling, and lack of collaborative evaluation of ecological suitability, and quality suitability. In this study, the maximum entropy model was used to analyze the ecological suitability of *P. notoginseng* under current and future climates. The multi-source chemical information of samples was collected to evaluate the uniformity between quality and ecology. The results showed that the current suitable habitat was mainly in southwest China. In the future climate scenarios, the high suitable habitat will be severely degraded. Modeling based on different regionalization could predict larger suitable habitat areas. The samples in the high suitable habitat had both quality suitability and ecological suitability, and the accumulation of chemical components had different responses to different environmental factors. Two-dimensional correlation spectroscopy combined with deep learning could achieve rapid identification of samples from different suitable habitats. In conclusion, global warming is not conducive to the distribution and spread of *P. notoginseng*. The high suitable habitat was conducive to the cultivation of high-quality medicinal materials. Actual regionalization modeling had more guiding significance for the selection of suitable habitats in a small area. The multi-regionalization modeling theory proposed in this study could provide a new perspective for the ecological suitability study of similar medicinal plants. The results provided a reference for the introduction and cultivation, and lay the foundation for the scientific and standardized production of high-quality *P. notoginseng*.

Keywords: *Panax notoginseng*, maximum entropy model, ecological suitability, quality suitability, multi-regionalization modeling

OPEN ACCESS

Edited by:
Yu Xue,
Nanjing University of Information Science and Technology, China

Reviewed by:
Rufeng Wang,
Shanghai University of Traditional Chinese Medicine, China
Jin Qi,
Nanjing University of Posts and Telecommunications, China

****Correspondence:***
YuanZhong Wang
boletus@126.com
ZhiTian Zuo
zzhitian@126.com

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

Received: *03 December 2021*
Accepted: *31 March 2022*
Published: *28 April 2022*

Citation:
Yue J, Li Z, Zuo Z and Wang Y (2022) Evaluation of Ecological Suitability and Quality Suitability of Panax notoginseng Under Multi-Regionalization Modeling Theory. Front. Plant Sci. 13:818376. doi: 10.3389/fpls.2022.818376

INTRODUCTION

Panax notoginseng (Burk.) F. H. Chen is a perennial herb of *Panax* genus, whose dry root and rhizome are valuable medicinal herbs, peculiar to China. Modern studies have proved that *P. notoginseng* mainly contains saponins, flavonoids, volatile oils, sterols, organic acids, polysaccharides, amino acids, and other chemical components (Wang et al., 2006, 2016, 2020; Li et al., 2013; Liu et al., 2020), and has anti-inflammatory, anti-tumor, protecting the cardiovascular system, improving the immunity and blood circulation and inhibiting the metastasis of liver tumor cells and other pharmacological effects (He et al., 2012; Zhou et al., 2014; Xie et al., 2018; My et al., 2021). Due to the uniqueness of the growth environment, *P. notoginseng* has become the dominant medicinal material in China. As the main raw material of traditional Chinese medicine preparation, decoction pieces, and Chinese patent medicine, it is also an important pillar for the development of the regional economy, which is of vital importance to the development of the traditional Chinese medicine industry and even the whole health industry (Cui et al., 2014). However, *P. notoginseng* is a typical ecologically fragile shade plant with a narrow distribution area. Coupled with problems, such as continuous cropping obstacles, diseases, insect pests, and other problems, the origin of *P. notoginseng* in Wenshan in Yunnan faces a planting bottleneck (Ou et al., 2021). In addition, there are reports of deforestation to plant *P. notoginseng* in 14 prefectures and cities in Yunnan, which poses varying degrees of harm to forest resources and the ecological environment (YUNNAN.CN, 2019). The contradiction between the protection of forest resources and the standardized development of *P. notoginseng* planting industry is becoming increasingly prominent. Accordingly, there is an urgent need for the introduction and expansion of new planting areas.

Panax notoginseng is an economic medicinal plant that has been cultivated and domesticated earlier. It has been artificially cultivated for more than 400 years, and its demand has been increasing sharply year after year (Wang et al., 2013). Traditionally, *P. notoginseng* in Wenshan has the best quality and effect. In order to meet the market demand, new planting areas have been developed in Kunming, Qujing, Honghe, and other regions. The quality differences in different regions reflect that the growing development of *P. notoginseng* is affected by natural ecological factors (Tao and Wu, 2003). Therefore, blind introduction and expansion will seriously affect the rational distribution of *P. notoginseng* production, greatly weakening the authenticity of medicinal materials, resulting in a serious decline in the quality of medicinal materials. When planting, it is extremely important to determine the suitable area for medicinal materials. For thousands of years, the evaluation of high-quality medicinal materials has been limited to the relationship between the efficacy and place of production. The understanding of the relationship between medicinal materials and the environment is still vague and superficial, failing to reveal its intrinsic relevance. The research on the ecology of medicinal plants and the ecology of traditional Chinese medicine resources should be actively strengthened to realize the synergy between the ecological suitability and quality suitability of medicinal plants.

Existing studies have analyzed the ecological suitability and quality ecology of *P. notoginseng* to different degrees. The possible distribution of the ecological suitable area and the main ecological factors affecting the geographical distribution were discussed (Zhang H.Z., et al., 2016). The study indicated that the possible ecological suitable areas in China were distributed in Yunnan, Guangxi, Guangdong, Guizhou, Hainan, Sichuan, Fujian, and Chongqing, accounting for more than 70% of the world's optimum cultivation area. Yunnan and Guangxi were the most suitable cultivation areas. The United States, Brazil, Portugal, and other countries also have a small number of optimal cultivation areas (Meng et al., 2016). The main ecological factors affecting the geographical distribution of *P. notoginseng* are the precipitation of warmest quarter, temperature seasonality, altitude, isothermality, coefficient of variation of precipitation seasonality, the mean of monthly, and the precipitation of the driest month (Zhang Q. et al., 2016). Although current studies can provide a theoretical basis for the suitable regionalization of *P. notoginseng*, there are some limitations, such as incomplete investigation of species distribution, single regionalization modeling method, and lack of ecological and quality coordination research, which seriously restrict the reliability of research results. In addition, most regionalization studies aimed at suitability accounted for the majority, while few regionalization studies focused on the accumulation of secondary metabolites. The regionalization studies did not discuss whether the growth suitability and quality suitability of the suitable area was consistent, resulting in that the suitable area obtained had no practical value for *P. notoginseng* cultivation, land use, and social economy.

The present study conducted a more comprehensive investigation and collection of the actual distribution of *P. notoginseng*. At the theoretical level, the interaction between the environment and planting of *P. notoginseng* was analyzed to find the key ecological factors, and the maximum entropy (MaxEnt) model was used to explore the potentially suitable habitat and simulate the impact of climate change on the distribution of *P. notoginseng* in the future. At the practical level, modern analytical methods were used to integrate multi-source chemical information of *P. notoginseng* samples, analyze the ecological suitability regionalization and quality suitability regionalization collaboratively, and establish a rapid identification method of *P. notoginseng* from high-quality suitable habitat. It is proposed to combine theory and practice on the basis of current species distribution, reveal the geographical distribution pattern of *P. notoginseng* under current and future climate scenarios, and evaluate the consistency of ecological suitability regionalization and quality suitability regionalization. In this way, the quality of the source is guaranteed, and a scientific basis is provided for the selection of suitable planting regionalization for *P. notoginseng* under future climate change.

MATERIALS AND METHODS

Species Distribution Data

The distribution data of *P. notoginseng* were obtained by three ways: (1) From 2011 to 2018, the Institute of Medicinal Plants, Yunnan Academy of Agricultural Sciences conducted field surveys on the major distribution areas in Yunnan, Guangxi, and surrounding areas, recorded the longitude and latitude information of planting sites in detail, and compiled data sets of species distribution. (2) Species distribution information was obtained through literature review and online sharing platform of specimen digital information. Platforms include the Chinese Virtual Herbarium databases of National Plant Specimen Resource Center (CVH, https://www.cvh.ac.cn), the National Specimen Information Infrastructure (NSII, www.nsii.org.cn), and the Global Biodiversity Information Facility (GBIF, https://www.gbif.org). (3) Government news reports combined with Google satellite images conducted investigation and image analysis of greenhouses for planting *P. notoginseng* in Yunnan, Guangxi, Guizhou, Sichuan, and in Chongqing from 2020 to 2021, so as to supplement and improve the species distribution data set. Through the integration of the above data sets, 3,475 species distribution data were obtained.

The species distribution data were further filtered and screened. First, the items with duplicate origin information and incomplete latitude and longitude information were removed. Then, based on Google Earth and the latitude and longitude information of specimens, the entries with incorrect origin information were revised or deleted. Finally, in order to avoid the redundancy of distribution data during modeling, ArcGIS was used to buffer and analyze species distribution data, and repeated data within 1 km was deleted to ensure that there was only one species distribution point in each 1.0 × 1.0 km^2 grid. Through the above screening, 2,122 species distribution data (**Figure 1**) were finally reserved for MaxEnt modeling analysis. Details of the distribution data are shown in **Supplementary Table S1**.

Environmental Variables

The environment variables for modeling mainly include elevation data and bioclimatic variables (**Table 1**). All these variables (spatial resolution of 30 s and raster data of about 1.0 × 1.0 km^2) were downloaded from the WorldClim (https://www.worldclim.org/). The climate variables required to study current species distributions were obtained from the grid data generated by interpolating observations from global climate stations from 1970 to 2000. Future climate data were predicted by the medium-resolution Climate System Model (BBC-CSM2-MR) developed by the National Climate Center (Beijing). Two greenhouse gas concentration pathways (medium-low concentration scenario: SSP245 and high concentration scenario: SSP585) under two future periods (2050s: 2041–2060 and 2090s: 2081–2100) were selected to study. The precision of future climate data provided by WorldClim is 2.5 min. In order to facilitate modeling, the data of spatial resolution was unified at 30 s by using ArcGIS raster resampling.

In order to avoid over-fitting of modeling results due to multicollinearity among environmental variables, elevation and bioclimatic variables corresponding to species distribution data were extracted for Spearman's correlation analysis. Combined with the modeling results, when the variables are significantly correlated and the correlation coefficient is >0.8, the variables with small modeling contributions are eliminated. All environment variables were converted to ASCII format before modeling. The final variables used for modeling are shown in **Table 1**.

Model Establishment and Evaluation

Regionalization of Modeling

Regionalization, which means the division of regions, is a way for people to extract spatial information according to a specific purpose, and to realize the understanding of the natural or the social environment through the region-based classification and merger (Margules and Pressey, 2000; Margules et al., 2002). The regionalization of traditional Chinese medicine resources is to study the spatial differentiation law of traditional Chinese medicine resources and their regional systems, and to divide them into regions according to this spatial consistency and difference law (Chinese Medical Company, 1995; Zhu et al., 2017). The traditional method of modeling based on all distribution points (i.e., overall regionalization) may have poor representativeness and ignore the internal relationship between the quality of medicinal materials and the regional ecological environment. As a result, the model results are only of theoretical significance and lack of practical guiding significance combined with actual regionalization. Through the comprehensive utilization of more reliable and reasonable species regionalization results, the collaborative modeling of ecological suitability regionalization and production suitability regionalization can greatly improve the practicability of predicted suitable habitat for *P. notoginseng*. Therefore, in addition to modeling for overall regionalization, this study also divided all distribution sites into the following regionalization according to the topographic characteristics of the distribution region, the actual planting situation, the authenticity of planting region, the traditional production area, the new production area, and other restrictions: (1) Hengduan Mountains, (2) Yunnan-Guizhou Plateau, (3) Yunnan + Guangxi, (4) Sichuan Basin, (5) Sichuan + Chongqing + Guizhou, and (6) Sichuan + Chongqing + Guizhou + Guangxi + Wenshan. It is planned to compare and predict the suitable habitat areas of *P. notoginseng* through different regionalization modeling.

Parameter Setting and Establishment of Model

MaxEnt modeling is one of the most widely used algorithms in the recent years (Phillips et al., 2006). This method has been widely used in the prediction of geographic distribution of medicinal plant species with high simulation accuracy (Guisan and Zimmermann, 2000; Elith et al., 2011). In this study, species distribution data in CSV format and environment variable data in ASCII format were imported into MaxEnt software (Version 3.4.4). About 75% of distribution data were set as training set and the remaining 25% as test set. Maximum iterations were set to 5,000, and the remaining parameters were default values. Model prediction results were saved as ASCII files.

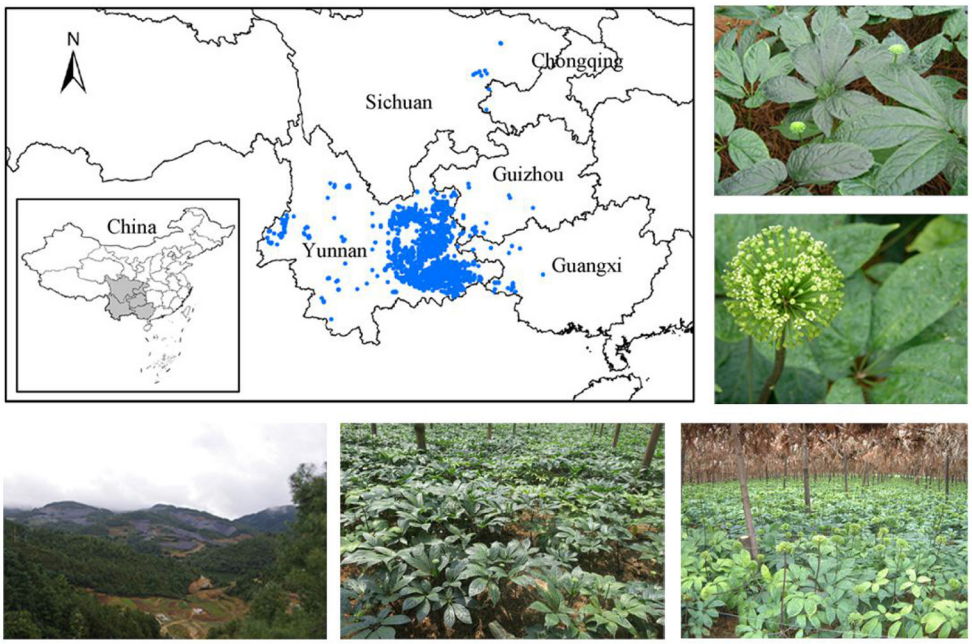


FIGURE 1 | The picture information on 2,122 distribution records, habitats, and samples of *Panax notoginseng*.

TABLE 1 | The environment variables information.

Type	Abbreviation	Environment variables	Unit
Climate factor	Bio01	Annual mean temperature	°C
	Bio02*	Mean diurnal range	°C
	Bio03*	Isothermality	/
	Bio04*	Temperature seasonality	°C
	Bio05	Max temperature of warmest month	°C
	Bio06	Min temperature of coldest month	°C
	Bio07*	Temperature annual range	°C
	Bio08	Mean temperature of wettest quarter	°C
	Bio09	Mean temperature of driest quarter	°C
	Bio10	Mean temperature of warmest quarter	°C
	Bio11	Mean temperature of coldest quarter	°C
	Bio12	Annual precipitation	mm
	Bio13*	Precipitation of wettest month	mm
	Bio14*	Precipitation of driest month	mm
	Bio15*	Precipitation seasonality	/
	Bio16	Precipitation of wettest quarter	mm
	Bio17*	Precipitation of driest quarter	mm
	Bio18*	Precipitation of warmest quarter	mm
	Bio19*	Precipitation of coldest quarter	mm
Topographic factor	Ele*	Elevation	m

Variables with an asterisk (*) were used for the final modeling.

The area under the curve (AUC) value of the area under the receiver operating characteristics (ROC) curve was used to evaluate the accuracy of the model results. The AUC value ranges from 0 to 1, and the closer the value is to 1, the more reliable the prediction results are (Hanley and McNeil, 1982; Zweig and Campbell, 1993). The index of important variables affecting the species distribution was determined by Jackknife test combined with the cumulative

contribution rate of modeling environmental variables (Phillips, 2005).

The ASCII files obtained by MaxEnt operation were imported into ArcGIS and converted into raster data. Raster data were reclassified according to habitat suitability index (HSI). The suitability was divided into four categories according to the difference of suitability index: high suitability (HIS $\geq$ 0.6), moderate suitability (0.4 $\leq$ HIS $<$ 0.6), low suitability (0.2 $\leq$ HIS $<$ 0.4), and no suitability (HIS $<$ 0.2) (Shen et al., 2021). The areas of various suitable habitat were calculated by Zonal Statistics of ArcGIS.

Post-processing of Model Results

The spatial extent of suitable habitat of *P. notoginseng* was modeled based on the current climate data. Since the overall regionalization is more representative of the actual distribution of *P. notoginseng*, the modeling projections of two future periods (2050s and 2090s) and two emission scenarios (SSP245 and SSP585) were performed based on the overall regionalization. Then, the HIS for current and future suitable habitats were calculated based on the results of MaxEnt software (probability of species presence). According to the suitability index, the habitat suitability of *P. notoginseng* was divided into high suitability, moderate suitability, low suitability and no suitability.

The changes of current and future suitable habitats were evaluated by calculating the suitable habitat change rate (SHCR) of high suitability, moderate suitability, and low suitability.

$$SHCR = (AF - AC)/AC.$$

Ecological Quality Suitability Analysis of *P. notoginseng*

Sample Collection

At present, the cultivation area of *P. notoginseng* is mainly concentrated in Yunnan, China. In addition, according to the survey results of actual distribution sites, there are more distribution sites in the five regions of Wenshan, Honghe, Yuxi, Kunming, and Qujing in Yunnan Province. Therefore, considering the actual planting situation, the economic value of production and the prediction results of suitable habitat for the Maxent model, these five regions were selected for sampling, and a total of 455 samples of *P. notoginseng* were collected. The picture information of some habitats, collection sites, and samples are shown in **Figure 1**. All samples were identified as *P. notoginseng* (Burk.) F. H. Chen by Professor JinYu Zhang, Institute of Medicinal Plants, Yunnan Academy of Agricultural Sciences. After sampling, the roots of fresh samples were taken, cleaned, and dried in a constant temperature oven at 50°C. Then all samples were crushed with a grinder, screened through 80 mesh, and stored in ziplock bags to avoid light for subsequent analysis. The number of samples used for content analysis and spectral analysis is shown in **Supplementary Table S1**.

Content Determination

In this study, the contents of notoginsenoside R_1, ginsenoside Rg_1, ginsenoside Re, and ginsenoside Rb_1 were determined by high-performance liquid chromatography (HPLC) analysis. The contents of total flavonoids and total polysaccharide were determined by UV-visible spectrophotometry. Chromatographic analyses were performed with LC-10 ATvp liquid chromatograph system with a diode array detector. binary pump, manual sampler, and CLASS-VP workstation. The chromatographic separation was achieved using an Inertsil ODS-3 column (4.6 $\times$ 150 mm, 5 μm). Standards of notoginsenoside R_1, ginsenoside Rg_1, ginsenoside Re, ginsenoside Rb_1, and rutin were purchased from the Chinese National Institute for Food and Drug Control (Beijing, China). Anhydrous glucose was purchased from Sigma-Aldrich. HPLC grade acetonitrile and methanol were purchased from Thermo Fisher Scientific (Fair Lawn, NJ, USA). Other chemicals and reagents were of analytical grade.

Determination of four kinds of saponins, such as Notoginsenoside R_1, ginsenoside Rg_1, ginsenoside Re, and ginsenoside Rb_1 were precisely measured at 2.00, 3.00, 2.20, and 2.40 mg, respectively, and dissolved in 1 mL of methanol solution. The solution was diluted 7 times according to the gradient of 60% to obtain a total of 8 mass concentration gradients of the reference solution. Sample powder was weighed accurately to 100 $\pm$ 0.30 mg and extracted with 2.0 mL of methanol by an ultrasound-assisted method for 60 min at ambient temperature. The methanol loss was supplemented after ultrasound, and the extract solution was filtered using a 0.22 μm membrane filter, which was the solution to be tested. The mobile phase and gradient elution procedures refer to HPLC detection of *P. notoginsen*g in the Chinese Pharmacopeia (I division, 2015 edition) (National Pharmacopoeia Committee, 2015). The mobile phase consisted of ultrapure water (A) and acetonitrile (B). The gradient elution sequence was conducted as follows: 0–12 min (19% B), 12–60 min (19–36% B). The column temperature was 30°C. The flow rate was kept at 1 mL/min and the injection volume was 10 μL. Detective wavelengths were set at 203 nm (Ma, 2014).

Determination of total flavonoid: Rutin was weighed accurately to 1.20 mg, dissolved in 8 mL of 70% ethanol solution, and then added 0.1 mol·L^{-1} of $AlCl_3$ ethanol solution (chromogenic agent) to 10 mL. The mother liquor of standard substance was diluted 6 times according to the equal gradient of 60% to obtain the standard substance solution with seven concentration gradients. Sample powder of 50 $\pm$ 0.2 mg were accurately weighed, and 8 mL of 70% ethanol solution was added for ultrasonic extraction for 1 h. Then, 6 mL of extracts were absorbed, and color reagent was added to 8 mL. Color developing time was 20 min, and scanning range was 190–600 nm. The background was scanned using a blank solution with chromogenic agent before collecting the sample spectra (Sun et al., 2018).

Determination of total polysaccharides: Anhydrous glucose was accurately weighed to be 30.50 mg, and the volume was adjusted to 100 mL with ultrapure water. The sample powder of 500.0 $\pm$ 1.0 mg was accurately weighed, and the volume was fixed to 50 mL with ultrapure water. Ultrasonic extraction was performed at 60°C water bath for 2 h, and the lost solution was supplemented after completion. The extract was quickly filtered by a quantitative filter paper. After the filtrate was cooled to

room temperature, 20 mL of filtrate was absorbed into a 50 mL centrifuge tube. After heating, the filtrate was volatilized to 10 mL, and then 100% ethanol was added to the filtrate at a constant volume of 50 mL. After blending, the filtrate was left to settle for 12 h. After precipitation, the solution was centrifuged at a speed of 9,000 r/min for 10 min, and the supernatant was extracted after centrifugation. The precipitation was washed with 30 mL of 80% ethanol solution and centrifuged at a speed of 9,000 r/min for 5 min. The supernatant was extracted and washed twice. The precipitation was added with ultra-pure water at a constant volume of 40 mL, and then heated in boiling water bath for 30 min after uniform oscillation, which was the solution to be tested. The glucose standard solution of 0.05, 0.1, 0.2, 0.4, 0.6, 0.8, and 1.0 mL was accurately pipetted into test tubes, and water is added to 2 mL, respectively. About 1.0 mL of the sample solution was precisely pipetted into a test tube, water was added to 2 mL, and finally, 1 mL of 5% phenol solution was added. After shaking, 4 mL of concentrated sulfuric acid was slowly added along the wall of the test tube, shaken quickly, and heated in a boiling water bath for 30 min. The solution was cooled to room temperature, and the UV scan range was set to 350–600 nm. The background was scanned using a blank solution with chromogenic agent before collecting the sample spectra (Li et al., 2017).

Fourier Transform Mid-infrared and Near Infrared Spectra Acquisition

The Fourier transform mid-infrared spectra (FT-MIR) of each samples were performed by the Fourier transform infrared spectrometer (Perkin Elmer, Norwalk, CT, USA) equipped with a deuterated triglycine sulfate (DTGS) detector. The sample powder (1.2 ± 0.2 mg) and KBr powder (15.0 ± 1.0 mg) were accurately weighed and mixed in an agate mortar. After grinding, the mixture was poured into a mold and pressed to make uniform flakes. Each sample tablet was tested to obtain FT-MIR spectra that was recorded in the region of 4,000–650 cm^{-1} with a resolution of 4 cm^{-1}, and a total of 16 scans were performed. Two scans were repeated for all samples to obtain an average spectrum. Prior to sample scanning, the background was scanned using blank KBr tablet to remove carbon dioxide and water interference.

The near infrared spectra (NIR) of sample powder were collected with an Antaris II spectrometer (Thermo Fisher Scientific, USA) equipped with an integrating sphere diffused reflection mode and Result 2.1 software. The wavenumber region was set to 10,000–4,000 cm^{-1}, and the resolution was set to 8 cm^{-1}. About 64 scans were accumulated for a single sample, and each sample was repeated for three times to obtain the average spectrum. The spectral record after acquisition is the logarithm of reciprocal reflectance.

Analysis of Content Data

To investigate whether *P. notoginseng* samples with ecological suitability (these samples are usually from highly suitable habitat areas) has quality suitability, the contents of notoginsenoside R_1, ginsenoside Rg_1, ginsenoside Re, ginsenoside Rb_1, total flavonoids, and total polysaccharides in these samples were compared and analyzed. In addition, we drew the boxplot for intuitively displaying the content data distribution characteristics of *P. notoginseng* samples from Wenshan, Honghe, Yuxi, Kunming, and Qujing.

Principal component analysis (PCA) is an unsupervised pattern recognition method, which transforms multiple related variables into a group of unrelated new variables by dimensionality reduction, while preserving the original sample information as much as possible (Abdi and Williams, 2010). In this study, simca software was used to conduct PCA on the content data and 20 environmental variables of five regions, in order to explore the similarity and differentiation of environmental variables, and the correlation between environmental variables and content.

Discriminant Evaluation Based on Deep Learning

Deep learning is the main research method adopted in the development of artificial intelligence at the present stage, which has unique advantages in image classification and object recognition (LeCun et al., 2015). The spectroscopy technology has the advantages of fast, lossless, and convenient. In particular, a two-dimensional correlation spectroscopy (2DCOS), which has the advantages of multi-discipline, can greatly improve the resolution of spectra through the increase of dimension, and enrich the information carried by the spectra (Noda, 1989). Combining it with deep learning for the identification of medicinal plants can give full play to the advantages of the two technologies and greatly improve the efficiency of identification analysis. In this study, a deep learning method combined with 2DCOS was developed to rapidly identify samples from different suitable habitat areas.

Generalized 2DCOS is an effective method to improve the spectral resolution and solve spectral overlap problem by designing interference variables, which is obtained by discrete generalized two-dimensional correlation spectral algorithm. Its dynamic spectra are expressed as S, and the expression is provided in Equation (1), where v is the variable and t is the external disturbance (Noda, 1993).

$$S(v) = \begin{bmatrix} y(v, t_1) \\ y(v, t_2) \\ y(v, t_3) \\ \cdot \\ \cdot \\ \cdot \\ y(v, t_m) \end{bmatrix} \quad (1)$$

Studies have shown that synchronous 2DCOS is more suitable for the identification of medicinal plants (Yue et al., 2021). Therefore, we combined it with deep learning to identify samples from different suitable areas. The synchronous spectral intensity is, and its expression is provided in Equation (2).

$$\Phi(v_1, v_2) = \frac{1}{m-1} S(v_1)^T \cdot S(v_2) \quad (2)$$

Matlab2017b software was used to automatically generate synchronous 2DCOS images. All images were stored in the

corresponding folder in the JPEG image format of 64 × 64-pixel size, which is used to build the ResNet model. All data sets were divided into training set (60%), test set (30%), and external validation set (10%) using the Kennard–Stone algorithm.

In this study, a 12-layer ResNet was established with a weight attenuation coefficient λ of 0.0001 and a learning rate of 0.01. **Supplementary Table S2** showed the ResNet network parameter configuration. The model structure is shown in **Supplementary Figure S1**, where the input data is synchronous 2DCOS images. The identification flow chart of ResNet is shown in **Supplementary Figure S2**. The training set was used to train the model. The Stochastic Gradient Descent (SGD) method was adopted to search for optimal parameters and minimize the loss function value, so as to obtain the optimal model. The test set was used to verify whether the performance of the final model was optimal. The external validation set was used to verify the generalization ability of the model.

RESULTS AND DISCUSSION

Model Evaluation and Important Environmental Variables

The AUC value can be used to evaluate the credibility and accuracy of the model results, and its value is between −1 and 1. For the AUC value, $0.60 < \text{AUC} \leq 0.70$ indicates that the model effect is fair, the value $0.70 < \text{AUC} \leq 0.80$ indicates that the model effect is good, the value $0.80 < \text{AUC} \leq 0.90$ indicates that the model effect is very good, and the value $0.90 < \text{AUC} \leq 1.00$ indicates that the model effect is excellent (Swets, 1988). The AUC values of the prediction results of the seven models were all higher than 0.92, indicating that all models had excellent fitting effects and high prediction accuracy.

Eleven climate factors were selected as indicators to simulate the suitable habitat of *P. notoginseng*, and the main climate factors affecting the distribution of *P. notoginseng* were obtained. **Supplementary Table S3** shows AUC values and important environmental variables of the seven regionalization models under current climate conditions. The results of important environmental variables were obtained from the comprehensive analysis of contribution analysis of variables and the jackknife test of variable importance, which were listed in the descending order of importance as follows. The variables that had a higher contribution to the overall regionalization model were Bio04, Bio07, Ele, and Bio17, and their cumulative contribution rate was 95.32%. For the Hengduan Mountain regionalization model, Bio04, Bio03, Bio14, Bio17, and Ele had a high contribution rate to the potential distribution of *P. notoginseng* with a cumulative contribution rate of 94.08%. The variables of Bio04, Ele, Bio02, Bio17, and Bio14 contributed significantly to the potential distribution of *P. notoginseng* in Yunnan-Guizhou Plateau regionalization model, with a cumulative contribution rate of 93.40%. For Yunnan + Guangxi regionalization model, Bio04, Bio07, and Bio17 contributed significantly to the potential distribution, with a cumulative contribution rate of 96.48%. The variables of Bio04, Ele, Bio03, and Bio19 had a high contribution rate to the potential distribution in the Sichuan Basin regionalization model, with a cumulative contribution rate of 100.00%. Bio18, Bio14, Bio15, Bio02, Ele, and Bio19 had a high contribution rate to the potential distribution of *P. notoginseng* in the Sichuan + Guizhou regionalization model, with a cumulative contribution rate of 93.84%. For Sichuan + Guizhou + Guangxi + Wenshan regionalization model, Bio17, Bio04, Ele, Bio13, Bio15, and Bio07 contributed significantly to the potential distribution of *P. notoginseng* with a cumulative contribution rate of 99.53%.

In summary, there are certain differences in the important environmental variables affecting the distribution of suitable habitats in different regions, which has reference significance for the introduction and cultivation of *P. notoginseng* in different regions. On the whole, the main ecological factors affecting the spatial distribution of *P. notoginseng* are Bio04, Ele, and Bio17. According to relevant literature reports, the elevation of the most suitable planting area of *P. notoginseng* is about 1,400–1,800 m, and the incidence of black spot will increase with the higher elevation beyond the suitable range (Dong et al., 2016). In addition, precipitation and temperature have a vital influence on the growth and distribution. *P. notoginseng* prefers humidity and warmth, but excessive high temperatures and excessive precipitation during the growth period will cause diseases to occur, and too little will not meet the growth needs (Zhang Q. et al., 2016). Therefore, special attention should be paid to the reasonable control of moisture and temperature during the cultivation and introduction process in different regions. From the similarity of important environmental variables, Models 2, 3, 6, and 7 are relatively similar, while Models 4 and 5 are significantly different. The results showed that there was a certain correlation between the existing ecological environment of different regionalization, which could provide a certain reference for the division of real regions, so that the regionalization results could be truly used in the production of *P. notoginseng*. In addition, similar climate is a prerequisite for introduction and cultivation, and it is more likely to expand *P. notoginseng* successfully in regionalization with high similarity to traditional producing areas.

Current Distribution of *P. notoginseng* Under Overall Regionalization Model

The AUC values of training set and test set data of MaxEnt model under the overall regionalization were 0.927 (**Supplementary Table S3**), indicating that the model has excellent fitting effect and high prediction accuracy. Based on MaxEnt prediction results, the potential distribution of *P. notoginseng* under the overall regionalization model was divided and visualized, as shown in **Figure 2**. The suitable habitat distribution of *P. notoginseng* under current climatic conditions was mainly in the southwest of China. Highly suitable habitat areas were concentrated in Yunnan Province, presenting a horizontal distribution with broad areas in the east and sparse areas in the west. In the western margin of Guangxi, the southwest margin of Guizhou, and the southern margin of Sichuan, there are highly suitable habitat areas with small area. The high suitable habitat in Yunnan were mainly distributed in

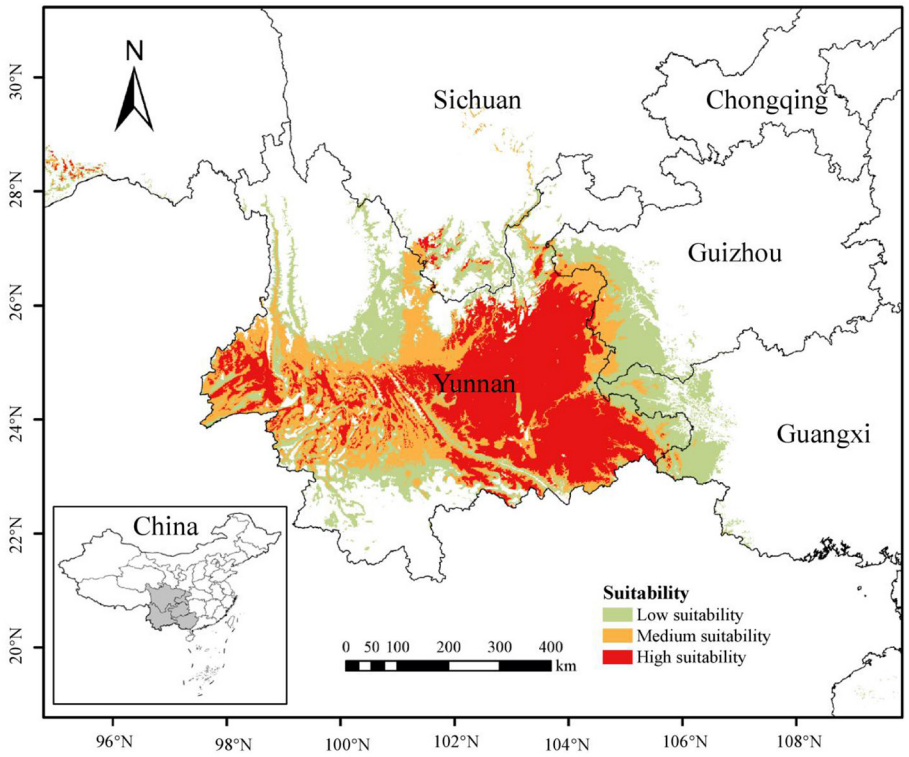


FIGURE 2 | The potential distribution of *Panax notoginseng* under current climate based on the overall regionalization model.

the central and southeast Yunnan, including Wenshan, Honghe, Qujing, Kunming, Yuxi, and Chuxiong. It is also distributed in some prefectures and cities of west and south Yunnan, such as Baoshan, Lincang, Dehong, and Pu'er. The moderate and low suitable habitat mainly include the central part of Hengduan Mountain region, the northwest part of Yunnan-Guizhou Plateau, and the borderlands of Yunnan, Guizhou, Guangxi, and Sichuan Province. Currently, the total area of suitable habitat in China is 27.85×10^4 km^2, of which the area of high suitable habitat is 10.01×10^4 km^2, the area of moderate suitable habitat is 8.44×10^4 km^2, and the area of low suitable habitat is 9.40×10^4 km^2 (**Supplementary Table S4**). The area of high suitable habitat in Yunnan is 9.74×10^4 km^2, accounting for 97.30% of that in China.

The model results show that the potential distribution area of *P. notoginseng* is relatively limited, which is consistent with the actual distribution of *P. notoginseng*. The main reason for this result is that *P. notoginseng* has special requirements on climate, soil, vegetation, and other growth environment (Zhang H.Z., et al., 2016). However, due to the rapid increase in the demand for *P. notoginseng* and the shortage of suitable land resources, the planting range has been expanded to many Prefectures and cities of Yunnan, Guangxi, Guizhou, Sichuan, and other places (Li et al., 2018). From the actual distribution and model prediction results, we can find that although there are suitable habitats in the northwestern Yunnan, *P. notoginseng* has not been planted on a large scale as in the southeastern Yunnan. We speculate that the accumulation of temperature and humidity in northwestern Yunnan is lower than that in the southeastern Yunnan since the barrier of Ailao Mountain is not conducive for the growth of *P. notoginseng*. In addition, the terrain in the northwestern Yunnan is steep. In terms of regional topography and geomorphology, this may also be one of the reasons why *P. notoginseng* has not been planted in large areas in the northwestern Yunnan. On the contrary, why Sichuan, as a new production area, can successfully plant *P. notoginseng* is also a question worthy of discussion and of great practical significance for the expansion of *P. notoginseng*. Judging from the distribution sites in Sichuan, almost all the planting sites are close to the river, which provides a guarantee for the important humidity conditions required for the growth of *P. notoginseng*. We speculate that this is one of the factors influencing its successful planting. Second, the accumulated temperature of Sichuan is similar to that of Yunnan, which may also be an important influencing factor. In future, the production of *P. notoginseng* will face the problem of further expansion of the production area and re-selection of the location. The current overall regionalization modeling is a general prediction of the suitable habitat, which has certain significance for actual planting. However, Chinese medicinal materials have spatial consistency and differences in climate, topography, landform, vegetation, soil, and other aspects according to natural conditions. Discussions on the suitability

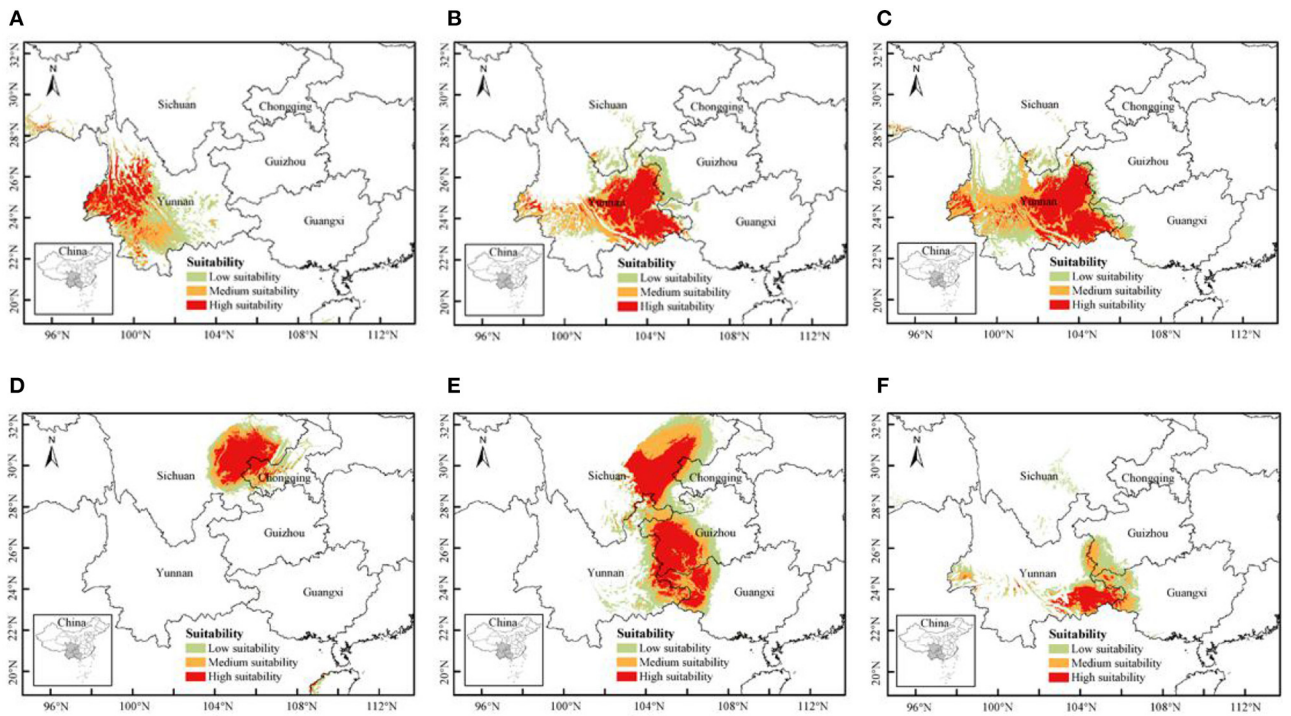


FIGURE 3 | The potential distribution of *Panax notoginseng* under current climate based on different regionalization model. **(A)** Hengduan Mountain regionalization, **(B)** Yunnan-Guizhou Plateau regionalization, **(C)** Yunnan + Guangxi regionalization, **(D)** Sichuan Basin regionalization, **(E)** Sichuan + Chongqing + Guizhou regionalization, and **(F)** Sichuan + Chongqing + Guizhou + Guangxi + Wenshan regionalization.

of Chinese medicinal materials based on regionalization may have greater reference value for actual production, planting, and quality. Therefore, it is necessary to explore suitable habitats for *P. notoginseng* under different regionalization according to local conditions.

Current Distribution of *P. notoginseng* Under Different Regionalization Models

MaxEnt model was used to predict the current potential distribution areas of *P. notoginseng* in different regionalization. The AUC values of the training set and test set data of the six models under different regionalization were all >0.920 (**Supplementary Table S3**), indicating that the model fitting effect was good and the prediction accuracy was high. The suitable habitat area data of the whole country and five provinces are shown in **Supplementary Table S4**, and the intuitive comparison chart is shown in **Supplementary Figure S3**. The distribution of suitable habitats for *P. notoginseng* in different regionalization models is shown in **Figure 3**.

Figure 3A shows the distribution of suitable habitat of *P. notoginseng* based on the regionalization of Hengduan Mountain. Compared to the overall regionalization model, the high suitable habitats in Hengduan Mountain region shifted and expanded to the northwest, and the area of low suitable habitats in Yunnan increased. **Figure 3B** shows the distribution of suitable habitat of *P. notoginseng* based on the regionalization model of Yunnan-Guizhou Plateau. Compared to the overall regionalization model, the regional change of high suitable habitat in Yunnan-Guizhou Plateau region is not significant, but the area of moderate and low suitable habitat is significantly reduced. **Figure 3C** shows the distribution of suitable habitat of *P. notoginseng* based on the regionalization of Yunnan + Guangxi. The regional distribution of high suitable habitat is roughly consistent with the overall regionalization model.

Figure 3D shows the suitable habitat distribution of *P. notoginseng* based on the regionalization of Sichuan Basin. **Figure 3E** shows the suitable habitat distribution of *P. notoginseng* based on Sichuan + Guizhou regionalization. The results of high, moderate, and low suitable habitat predicted by the two regionalization models were significantly different from those of the overall regionalization model. From the **Figure 3** and **Supplementary Table S4**, we can see that the area of high suitable habitat increased significantly in Sichuan, Guizhou, and Guangxi. Based on the Sichuan Basin Regionalization model, the high suitable habitat area in Sichuan is 46 times more than that under the overall regionalization model. The area of high suitable habitat in Sichuan, Guangxi, and Guizhou based on the Sichuan + Guizhou regionalization model is 57 times, 91 times, and 104 times that of Sichuan, Guangxi, and Guizhou based on the overall regionalization model, respectively. **Figure 3F** shows the suitable habitat distribution of *P. notoginseng* based on the Sichuan + Guizhou + Guangxi + Wenshan regionalization model. This regionalization added Guangxi and Wenshan on the basis of Sichuan + Guizhou regionalization. The prediction

results of the model are greatly different from the results of model in **Figure 3E**. The high suitable habitat area is mainly distributed in Wenshan, with an area of only 1.71 × 10^4 km^2.

In general, compared to the overall regionalization model, the distribution of high suitable habitat in the Yunnan-Guizhou Plateau regionalization model and Yunnan + Guangxi regionalization model was consistent, while the other four regionalization models (A, D, E, and F) had significant changes in both area and location. This was caused by the differences of topographic, geomorphic, and climatic factors in different regionalization, which indicated that selecting the suitable habitat in these four divisions should refer to the prediction results of each regionalization rather than the overall regionalization. From models D, E, and F, we can see that Sichuan, Guangxi, and Guizhou should theoretically have larger suitable habitat areas. Since the overall regionalization model is more focused on traditional production areas with more distribution points, it conceals important environmental factors of new production areas, and underestimates the potential and suitability of planting *P. notoginseng* in Sichuan, Guangxi, and Guizhou. It is worth noting that the reason why the suitable area of the F model is so different from that of the E model may be due to the difference in the altitude. Therefore, for areas with large differences in topography and geomorphology, modeling according to different regionalization may be more practical in guiding the selection of suitable habitats in different regions than the overall regionalization modeling. In this study, modeling based on different regionalization can predict a larger area of suitable habitat and provide more choices for the introduction of *P. notoginseng*. However, a more reasonable and practical regionalization study is needed to provide a solid foundation for the suitable habitat analysis of *P. notoginseng*.

Future Distribution of *P. notoginseng* Under Overall Regionalization Model

Under the background of climate change, two climate scenarios of 2050s and 2090s were selected respectively. The optimum MaxEnt model was used to simulate the geographical distribution of *P. notoginseng*, and the suitable habitat distribution (**Figure 4**) and suitable habitat area (**Table 2**) under different climate scenarios in different periods was obtained. In order to evaluate the distribution changes of suitable habitat of *P. notoginseng* under current and future climate conditions, we also calculated the SHCR (**Table 2**). The results show that the total suitable habitat areas for low, moderate, and total areas increased under the four scenarios in the two periods. Among them, the areas of low, moderate, and total suitable habitats under the SSP585-2050s climate background increased the most, and their areas were 14.70 × 10^4 km^2, 11.07 × 10^4 km^2, and 33.96 × 10^4 km^2. Compared to the current climate scenario, the area of high suitable habitats has increased by 56.38, 31.16, and 21.94%, respectively. It is worth noting that under different climate scenarios, the increase of low suitable habitat is the most significant, and the SHCR is higher than 40%. The area of high suitable habitat decreased, and the area of high suitable habitat was 8.95 × 10^4 km^2, 8.58 × 10^4 km^2, 8.19 × 10^4 km^2, and 8.02 × 10^4 km^2 under the four scenarios, respectively. Compared to the current climate, the area of high suitable habitat decreased by 10.59, 14.29, 18.18, and 19.88%, respectively. In the scenario of SSP585-2090s, *P. notoginseng* has the smallest area of high suitable area in the whole country.

In this study, it was found that *P. notoginseng* has the widest distribution and the largest area of high suitable habitat under the current climate scenario. With the increase of time and emission intensity, the area of high suitable habitat decreased, and the emission intensity had a greater impact on the area of high suitable habitat than that of time. The higher the emission intensity in the same period, the more is the decrease in the area of high suitable habitat, indicating that the scenario of high concentration of CO_2 emission had a greater impact on the potential distribution area of *P. notoginseng*. The area of high suitable habitat of *P. notoginseng* decreased in China and other provinces, indicating that the high suitable habitat of *P. notoginseng* will be degraded seriously under the background of future climate change. Yunnan was the province with the largest distribution of high suitable habitat area under different climate scenarios in different periods. The total suitable habitat area increased significantly, while the high suitable habitat area decreased, indicating that the high suitable habitat was transforming to the moderate and low suitable habitat, and the moderate and low suitable habitat was expanding continuously. The transformation and expansion of suitable habitat in Sichuan was consistent with that in Yunnan under different climate scenarios in different periods, while the total suitable habitat area in Guangxi and Guizhou decreased, and the moderate and low suitable habitat area showed slight expansion and contraction. The suitable habitat areas in Yunnan and Sichuan under different climate scenarios in various periods will continue to expand in the future, and the phenomenon with the period and emissions intensity had no obvious positive correlation, indicating that Yunnan is still the main area for the cultivation and expansion of *P. notoginseng* both at present and in the future, and Sichuan is the potential area of expansion in the future. But it should be noted that the introduction and cultivation should be adapted to local conditions. A small-scale introduction should be carried out first, and then a large-scale promotion can be carried out after successful trial planting. In addition, ecological planting modes based on regional distribution have gradually developed, such as intercropping of medicine and grain, cropping rotation and under-forest planting (Kang et al., 2020), which can effectively alleviate land shortage caused by continuous cropping obstacles under traditional planting mode, provide high-quality medicinal materials, and promote the development of multi-dimensional cultivation mode. These cropping patterns were coupled with species interaction and habitat demand for growing environment. However, climate change will affect species richness and community stability. Therefore, in addition to the prediction of a single species in the study of suitable habitats for *P. notoginseng* in the future climate, it is also

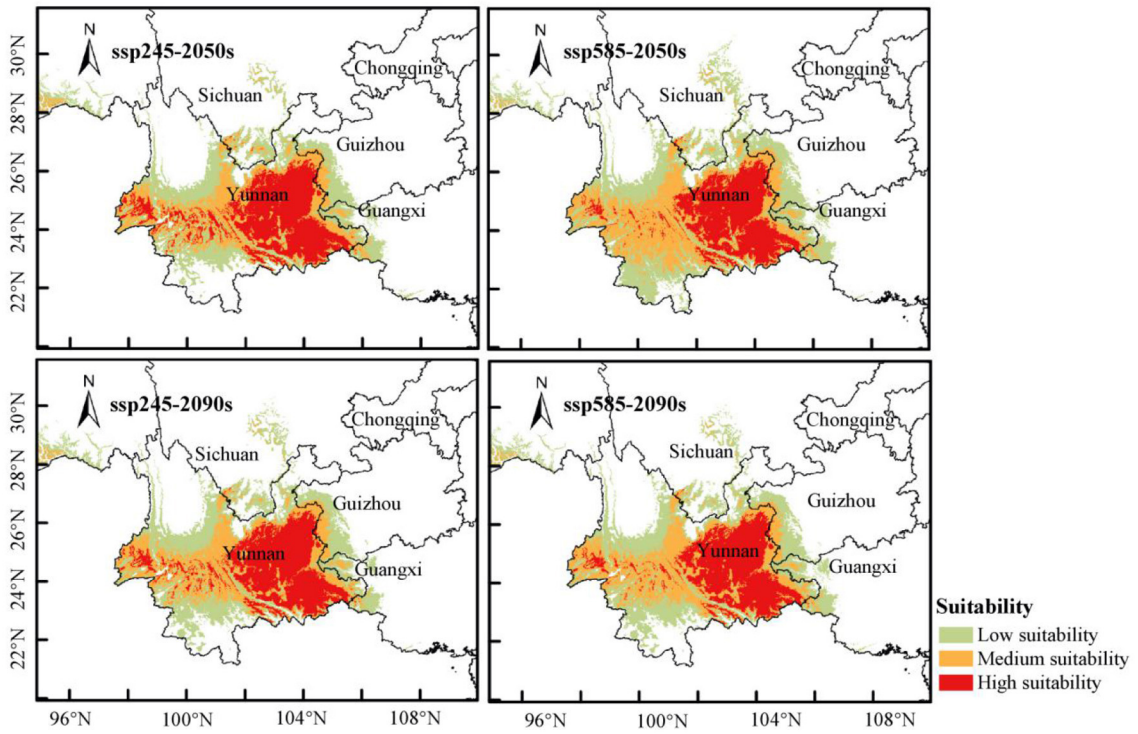


FIGURE 4 | Future potential distribution of *Panax notoginseng* under climate change scenarios SSP245 and SSP585.

TABLE 2 | The distribution area and suitable habitat change rate of *Panax notoginseng* under future climate scenarios.

Scenario	Area ($\times 10^4$ km^2)				SHCR (%)			
	L	M	H	T	L	M	H	T
ssp245-2050s	12.40	8.90	8.95	30.25	31.91	5.45	−10.59	8.62
ssp245-2090s	13.59	9.34	8.58	31.51	44.57	10.66	−14.29	13.14
ssp585-2050s	14.70	11.07	8.19	33.96	56.38	31.16	−18.18	21.94
ssp585-2090s	13.65	9.90	8.02	31.57	45.21	17.30	−19.88	13.36

L, Low suitable habitat; M, Moderate suitable habitat; H, High suitable habitat; T, Total suitable habitat; SHCR, Suitable habitat change rate.

necessary to coordinate the study of suitable habitats between regional communities.

Quality Suitability Analysis

Quality Evaluation of Samples in High Suitable Habitat

The contents of notoginsenoside R_1, ginsenoside Rg_1, ginsenoside Re, ginsenoside Rb_1, total flavonoids, and total polysaccharides of *P. notoginseng* samples from five high suitable habitats (Wenshan. Honghe, Yuxi, Kunming, and Qujing) were determined (**Supplementary Table S5**). **Supplementary Figure S4** is the boxplot of content data. As can be seen from the **Supplementary Figure S4** and **Supplementary Table S5**, the content of ginsenoside Rg_1 is the highest, followed by ginsenoside Rb_1, notoginsenoside R_1 is the third, and the content of ginsenoside Re is the lowest among the four saponins. Among the five regions, notoginsenoside R_1 and ginsenoside Rg_1 in Honghe, ginsenoside Re, and total flavonoids in Wenshan, ginsenoside Rb_1 in Yuxi, and total polysaccharides in Kunming were the highest. Taken together, the samples from Honghe had better quality than those from the other high suitable habitat according to the standard of "Pharmacopeia of People's Republic of China." We speculated that the difference in content was caused by the elevation of the collection sites, changeable climatic conditions, or other natural factors, which led to the genetic variation of the samples or the change of their own metabolites.

In general, the contents of notoginseng R_1, ginsenoside Rg_1, and ginsenoside Rb_1 in 50 samples of *P. notoginseng* were higher than the standard contents stipulated by Chinese Pharmacopeia of 2020 Edition (National Pharmacopoeia Committee, 2020), indicating that the *P. notoginseng* samples in the five high suitable habitats had high quality. It can be seen from the above results that the ecological suitability and quality suitability of *P.*

notoginseng samples in the high suitable habitat are consistent. However, this study also has some problems, such as insufficient sample size and coverage of sample collection points in different regions, so further research is needed.

PCA Analysis Based on Content Data and Environmental Factors

Principal Component Analysis is a multivariate statistical technique, which simplifies data structure by simplifying multiple variables into a few comprehensive variables with the idea of dimensionality reduction. PCA analysis was carried out on the content data of *P. notoginseng* and corresponding environmental variables of different sampling sites in five regions. This study revealed the relationship between environmental factors and seven active ingredients using principal components. **Supplementary Figure S5** is the scree plot of PCA, which is a graph of eigenvalues sorted from the largest to the smallest, used to determine the number of principal components. The first five principal components explained 92.3% of the variables cumulatively, of which 41.2% of the variables were explained by PC1. **Supplementary Figure S6** is a loading plot for PCA, showing the contribution rate of each variable to the principal component. The red dashed line on the graph above indicates the expected average contribution. For a given component, a variable with a contribution larger than this cutoff could be considered as important in contributing to the component. **Supplementary Figure S7** is a variable correlation plot colored by groups. It shows the relationships between all variables. It can be interpreted as follow: Positively correlated variables are grouped together. Negatively correlated variables are positioned on opposite sides of the plot origin (opposed quadrants). The distance between variables and the origin measures the quality of the variables on the factor map. Variables that are away from the origin are well-represented on the factor map.

Obviously, Bio12, Bio16, Bio18, and Bio13 have a positive effect on the accumulation of GB, GE, and NR, especially GB and GE. However, Bio07, Bio02, and Ele are not conducive to the accumulation of these three components, but they are conducive to the accumulation of TP. Bio14, Bio04, Bio19, and Bio17 are beneficial to the accumulation of TF, but not beneficial to the accumulation of NR and GR. There are many environmental variables that have a positive effect on GR accumulation, and all of them are distributed in the fourth quadrant. In particular, Bio05 has a significant effect on GR accumulation. The study results of Feng et al. (2006) on *P. notoginseng* showed that the precipitation in January and the annual temperature difference were the key factors affecting the total saponin of *P. notoginseng*. The precipitation was beneficial to the accumulation of flavonoid, but inhibited the accumulation of total saponin, polysaccharides, and dencichine in *P. notoginseng*. This is consistent with the correlation analysis results of this study. In addition, other related studies also showed that the region with small annual temperature difference was conducive to the accumulation of notoginseng saponin. Excessive precipitation would inhibit the accumulation of notoginseng saponin, while the overcast and rainy environment with more water was conducive to the accumulation of flavonoid, but not conducive to the accumulation of dencichine (Dong et al., 2003). Therefore, in the cultivation of *P. notoginseng*, appropriate artificial control of relevant environmental factors, such as light, humidity and so on, is conducive to the accumulation of important characteristic components of *P. notoginseng*.

Rapid Identification of Samples From Different Suitable Habitat

In this study, ResNet models were established based on FT-MIR and NIR 2DCOS images to identify *P. notoginseng* samples from different suitable areas. The discriminative ability of all models was evaluated by accuracy and loss values. External validation was used to judge and evaluate the strengths and weaknesses of the models to ensure the stability of the established model.

Figure 5 shows the accuracy curve, cross entropy, and external verification results of the established models. **Figure 5A** is the identification result based on the FT-MIR 2DCOS images. The results show that the accuracy of the training set is 99%, the test set is 99%, the external validation set is 100%, the number of epochs is 20, and the loss value is 0.053. **Figure 5B** is the identification result based on the NIR 2DCOS images. The results show that the accuracy of the training set is 100%, test set is 99%, the external validation set is 100%, the number of epochs is 20, and the loss value is 0.056.

According to the modeling results, ResNet models based on FT-MIR and NIR data can achieve rapid identification of *P. notoginseng* samples from five suitable areas. From the perspective of accuracy rate, lost value, and external verification results, the two models have almost no difference and have excellent discrimination performance with almost 100% accuracy. This result may be related to the learning performance of the deep learning model itself, which can extract features of different images for learning and finally realize the identification of target objects (Bargshady et al., 2020). The accuracy curves based on the training set and test set are highly consistent with the increasing trend of training rounds, which indicates that the model has no over-fitting risk and good robustness. The loss values of these models are close to zero, which verifies that the established models have good convergence effect.

In conclusion, deep learning combined with 2DCOS is an efficient and simple recognition method, which eliminates the need to manually extract sample data features and reduces the interference of human factors. As a means of identification, it is applied after ecological suitability evaluation and quality suitability evaluation, so as to realize the rapid identification and evaluation of *P. notoginseng* samples with both ecological suitability and quality suitability.

CONCLUSION

This study combined theory with practice to study the geographical distribution pattern of *P. notoginseng* under current and future climate scenarios and predicted the suitable habitat of *P. notoginseng* through overall regionalization model and different regionalization model. The ecological suitability and quality suitability of *P. notoginseng* samples were analyzed by integrating multi-source chemical information of samples from

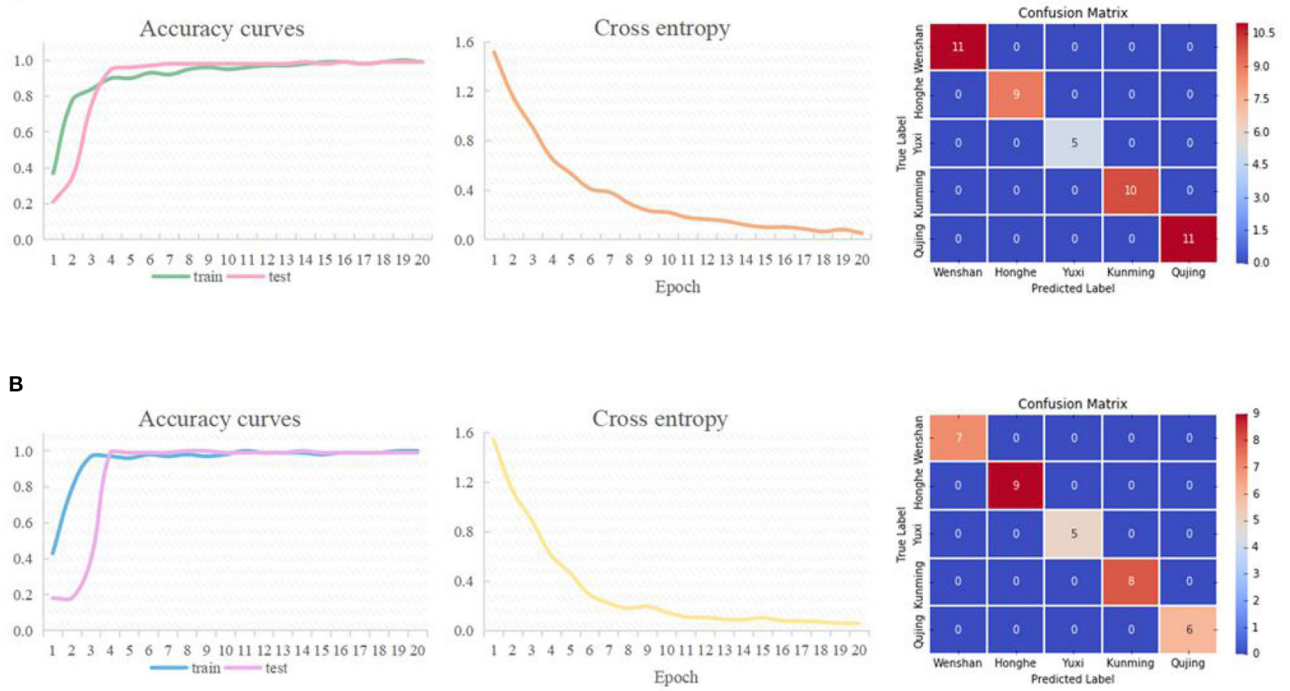


FIGURE 5 | The accuracy curves, cross entropy, and confusion matrix of ResNet models based on different spectral data. **(A)** FT-MIR, **(B)** NIR.

different suitable habitats with modern analytical methods, and the rapid identification method of high-quality samples from different suitable habitats was established. Under the overall regionalization model, *P. notoginseng* is mainly distributed in the southwest China, and its high suitable habitat areas are mainly concentrated in Yunnan, China. Under future climate scenarios, the high suitable habitat will be severely degraded, and greenhouse gas emissions will not be conducive to the distribution and diffusion of *P. notoginseng*. However, the suitable habitat areas in Yunnan and Sichuan are expanding continuously. Whether at present or in the future, Yunnan is still the main area for the cultivation and expansion of *P. notoginseng*, and Sichuan is a region with a great potential for the expansion of *P. notoginseng* in the future. Modeling based on different regionalization can more accurately predict suitable habitats in a smaller area. Especially in Sichuan, Guizhou, Guangxi, and other places, the suitable habitat under different regionalization model is much higher than that of the overall regionalization model, which can provide more choices for the introduction and cultivation of *P. notoginseng*. The analysis of multi-source chemical information shows that the samples of *P. notoginseng* in the high suitable habitat have higher quality, and the ecological suitability and quality suitability are consistent. In addition, changes in habitat will also affect the quality of medicinal materials. The accumulation of saponin, polysaccharide, and flavonoid in the samples is affected by various environmental factors. Proper regulation of relevant environmental factors during the planting process is conducive to the improvement of the important characteristic components of *P. notoginseng*. Deep learning combined with 2DCOS can realize the identification of samples in different suitable areas, providing an efficient and simple method for the evaluation of Traditional Chinese medicine. This study can guide the reasonable layout of production base, and the introduction and breeding, and promote the sustainable development of high-quality *P. notoginseng* industry.

DATA AVAILABILITY STATEMENT

The original contributions presented in the study are included in the article/**Supplementary Materials**, further inquiries can be directed to the corresponding authors.

AUTHOR CONTRIBUTIONS

JY: conceptualization, methodology, software, validation, formal analysis, and writing-original draft preparation. ZL: resources, validation, and writing-review and editing. ZZ: supervision, project administration, and funding acquisition. YW: resources, supervision, project administration, and funding acquisition. All authors have read and agreed to the published version of the manuscript.

FUNDING

This work was supposed by National Natural Science Foundation of China (Grant Number: 31760360),

Natural Science Foundation of Yunnan Province of China (Grant Number: 202101AT070260), and Special Program for the Major Science and Technology Projects of Yunnan Province (Grant Number: 202002AA100007).

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fpls.2022.818376/full#supplementary-material

REFERENCES

Abdi, H., and Williams, L. J. (2010). Principal component analysis. *Wiley Interdiscipl. Rev. Comput. Stat.* 2, 433–459. doi: 10.1002/wics.101

Bargshady, G., Zhou, X. J., Deo, R. C., Soar, J., Whittaker, F., and Wang, H. (2020). Enhanced deep learning algorithm development to detect pain intensity from facial expression images. *Expert Syst. Appl.* 149, 113305. doi: 10.1016/j.eswa.2020.113305

Chinese Medical Company (1995). *Chinese Material Medical Regionalization*. Beijing: Science Press.

Cui, X. M., Huang, L. Q., Guo, L. P., and Liu, D. H. (2014). Chinese sanqi industry status and development countermeasures. *China J. Chin. Mater. Med.* 39, 553–557. doi: 10.4268/cjcmm20140401

Dong, S. F., Li, W. M., Li, M. C., and Guan, M. F. (2016). Discussion on cultivation of *Panax notoginseng* in Honghe prefecture. *Agric. Technol.* 36, 111. doi: 10.11974/nyyjs.20160133092

Dong, T. T. X., Cui, X. M., Song, Z. H., Zhao, K. J., Ji, Z. N., Lo, C. K., et al. (2003). Chemical assessment of roots of *Panax notoginseng* in China: regional and seasonal variations in its active constituents. *J. Agric. Food Chem.* 51, 4617–4623. doi: 10.1021/jf034229k

Elith, J., Phillips, S. J., Hastie, T., Dudík, M., Chee, Y. E., and Yates, C. J. (2011). A statistical explanation of MaxEnt for ecologists. *Divers. Distrib.* 17, 43–57. doi: 10.1111/j.1472-4642.2010.00725.x

Feng, X. Q., Cui, X. M., Chen, Z. J., Zhang, Y. L., and Zhang, W. S. (2006). Analysis of correlation between effective components of burk (*Panax notoginseng*) and meteorological factors. *Chin. J. Agrometeor.* 1, 16–18.

Guisan, A., and Zimmermann, N. E. (2000). Predictive habitat distribution models in ecology. *Ecol. Model.* 135, 147–186. doi: 10.1016/S0304-3800(00)00354-9

Hanley, J. A., and McNeil, B. J. (1982). The meaning and use of the area under a receiver operating characteristic (ROC) curve. *Radiology* 143, 29–36. doi: 10.1148/radiology.143.1.7063747

He, N. W., Zhao, Y., Guo, L., Shang, J., and Yang, X. B. (2012). Antioxidant, antiproliferative, and pro-apoptotic activities of a aaponin extract derived from the roots of *Panax notoginseng* (Burk.) F.H. *Chen. J. Med. Food* 15, 350–359. doi: 10.1089/jmf.2011.1801

Kang, C. Z., Lu, C. G., Huang, L. Q., Wang, S., Wang, H. Y., Zhang, W. J., et al. (2020). Pattern of ecological planting for Chinese materia medica based on regional distribution. *China J. Chin. Mater. Med.* 45, 1982–1989. doi: 10.19540/j.cnki.cjcmm.20200302.103

LeCun, Y., Bengio, Y., and Hinton, G. (2015). Deep learning. *Nature* 521, 436–444. doi: 10.1038/nature14539

Li, S. P., Qiao, C. F., Chen, Y. W., Zhao, J., Cui, X. M., Zhang, Q. W., et al. (2013). A novel strategy with standardized reference extract qualification and single compound quantitative evaluation for quality control of *Panax notoginseng* used as a functional food. *J. Chromatogr. A* 1313, 302–307. doi: 10.1016/j.chroma.2013.07.025

Li, Y., Zhang, J., Xu, F. R., Liu, F., Wang, Y. Z., and Zhang, J. Y. (2017). Multiple attribute decision making analysis TOPSIS on quality evaluation study of *Panax notoginseng. Chin. Trad. Herb. Drugs* 48, 4764–4771. doi: 10.7501/j.issn.0253-2670.2017.22.027

Li, Y., Zhang, J. Y., and Wang, Y. Z. (2018). FT-MIR and NIR spectral data fusion: a synergetic strategy for the geographical traceability of *Panax notoginseng. Anal. Bioanal. Chem.* 410, 91–103. doi: 10.1007/s00216-017-0692-0

Liu, L., Xu, F. R., and Wang, Y. Z. (2020). Traditional uses, chemical diversity and biological activities of *Panax L.* (Araliaceae): a review. *J. Thnopharmacol.* 263, 112792. doi: 10.1016/j.jep.2020.112792

Ma, W. X. (2014). Simultaneous determination of ginsenoside Rb1 ginsenoside Rg1 and notoginsenoside R1 in *Panax notoginseng* by HPLC gradient elution. *China Pharmac.* 23, 43–44.

Margules, C. R., and Pressey, R. L. (2000). Systematic conservation planning. *Nature* 405, 243–253. doi: 10.1038/35012251

Margules, C. R., Pressey, R. L., and Williams, P. H. (2002). Representing biodiversity: data and procedures for identifying priority areas for conservation. *J. Biosci.* 27, 309–326. doi: 10.1007/BF02704962

Meng, X. X., Huang, L. F., Dong, L. L., Li, X. W., Wei, F. G., Chen, Z. J., et al. (2016). Analysis of global ecology of *Panax notoginseng* in suitability and quality. *Acta Pharm. Sinic.* 51, 1483–1493. doi: 10.16438/j.0513-4870.2016-0733

My, P. L. T., My, H. T. K., Phuong, N. T. X., Dat, T. D., Thanh, V. H., Nam, H. M., et al. (2021). Optimization of enzyme-assisted extraction of ginsenoside Rb1 from Vietnamese *Panax notoginseng* (BURK.) F.H. Chen roots and anticancer activity examination of the extract. *Sep. Sci. Technol.* 56, 1687–1698. doi: 10.1080/01496395.2020.1795676

National Pharmacopoeia Committee (2015). *Pharmacopoeia of People's Republic of China*. Beijing: China Medical and Technology Press.

National Pharmacopoeia Committee (2020). *Pharmacopoeia of People's Republic of China*. Beijing: China Medical and Technology Press.

Noda, I. (1989). Two-dimensional infrared spectroscopy. *J. Am. Chem. Soc.* 111, 8116–8118. doi: 10.1021/ja00203a008

Noda, I. (1993). Generalized two-dimensional correlation method applicable to infrared, raman, and other types of spectroscopy. *Appl. Spectrosc.* 47, 1329–1336. doi: 10.1366/0003702934067694

Ou, X. H., Cui, X. M., Zhu, D. W., Guo, L. P., Liu, D. H., and Yang, Y. (2021). Cultivation mode of *Panax notoginseng* causes NH4+ accumulation in planting soil. *Arch. Agron. Soil Sci.* 67, 960–973. doi: 10.1080/03650340.2020.1771314

Phillips, S. J. (2005). A brief tutorial on Maxent. *AT&T Res.* 190, 231–259.

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

Shen, T., Yu, H., and Wang, Y. Z. (2021). Assessing the impacts of climate change and habitat suitability on the distribution and quality of medicinal plant using multiple information integration: take Gentiana rigescens as an example. *Ecol. Indic.* 123, 107376. doi: 10.1016/j.ecolind.2021.107376

Sun, H., Jing, H., Hu, Q., Kang, P. D., Chen, J. F., He, J. L., et al. (2018). Infrared spectroscopy combined with Chemometrics for rapid determination of total flavonoids in *Dendrobium officinale. Spectrosc. Spect. Anal.* 38, 1702–1707. doi: 10.3964/j.issn.1000-0593(2018).06-1702-06

Swets, J. A. (1988). Measuring the accuracy of diagnostic systems. *Science* 240, 1285–1293. doi: 10.1126/science.3287615

Tao, S. H., and Wu, F. E. (2003). Effect of ecological environment on active constituents of medicinal plants. *Nat. Prod. Res. Dev.* 15, 174–177. doi: 10.16333/j.1001-6880.2003.02.024

Wang, C. Z., McEntee, E., Wicks, S., Wu, J. A., and Yuan, C. S. (2006). Phytochemical and analytical studies of *Panax notoginseng* (Burk.) F.H. Chen. *J. Nat. Med.* 60, 97–106. doi: 10.1007/s11418-005-0027-x

Wang, D., Koh, H. L., Hong, Y., Zhu, H. T., Xu, M., Zhang, Y. J., et al. (2013). Chemical and morphological variations of *Panax notoginseng* and their relationship. *Phytochemistry* 93, 88–95. doi: 10.1016/j.phytochem.2013.03.007

Wang, Q., Mu, R. F., Liu, X., Zhou, H. M., Xu, Y. H., Qin, W. Y., et al. (2020). Steaming changes the composition of saponins of *Panax notoginseng* (Burk.) F.H. Chen that function in treatment of hyperlipidemia and obesity. *J. Agric. Food Chem.* 68, 4865–4875. doi: 10.1021/acs.jafc.0c00746

Wang, T., Guo, R. X., Zhou, G. H., Zhou, X. D., Kou, Z. Z., Sui, F., et al. (2016). Traditional uses, botany, phytochemistry, pharmacology and toxicology

of *Panax notoginseng* (Burk.) F. H. Chen: a review. *J. thnopharmacol.* 188, 234–258. doi: 10.1016/j.jep.2016.05.005

Xie, W. J., Meng, X. B., Zhai, Y. D., Zhou, P., Ye, T. Y., Wang, Z., et al. (2018). *Panax notoginseng* saponins: a review of its mechanisms of antidepressant or anxiolytic effects and network analysis on phytochemistry and pharmacology. *Molecules* 23, 940. doi: 10.3390/molecules23040940

Yue, J. Q., Huang, H. Y., and Wang, Y. Z. (2021). A practical method superior to traditional spectral identification: two-dimensional correlation spectroscopy combined with deep learning to identify Paris species. *Microchem. J.* 160, 105731. doi: 10.1016/j.microc.2020.105731

YUNNAN.CN (2019). *Yunnan Province Will Launch a Three-Month Campaign Against the Deforestation of Panax notoginseng in 14 Cities*. Available online at: https://yn.yunnan.cn/system/2019/03/01/030214584.shtml

Zhang, H. Z., Liu, D. H., Zhang, D. K., Wang, Y. H., Li, G., Yan, G. L., et al. (2016). Quality assessment of *Panax notoginseng* from different regions through the analysis of marker chemicals, biological potency and ecological factors. *PLoS ONE* 11, e164384. doi: 10.1371/journal.pone.0164384

Zhang, Q., Zeng, F. L., Zhang, D. F., Xie, C. X., and Chen, S. L. (2016). Ecology suitability regions and ecological characteristics of *Panax notoginseng* (Burk.) F.H.Chen based on maximum entropy model. *Acta Pharm. Sinic.* 51, 1629–1637. doi: 10.16438/j.0513-4870.2016-0148

Zhou, N. N., Tang, Y., Keep, R. F., Ma, X. X., and Xiang, J. M. (2014). Antioxidative effects of *Panax notoginseng* saponins in brain cells. *Phytomedicine* 21, 1189–1195. doi: 10.1016/j.phymed.2014.05.004

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, 1–14. doi: 10.1038/srep42417

Zweig, M. H., and Campbell, G. (1993). Receiver-operating characteristic (ROC) plots: a fundamental evaluation tool in clinical medicine. *Clin. Chem.* 39, 561–577. doi: 10.1093/clinchem/39.4.561

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: 06 May 2022
doi: 10.3389/fpls.2022.915811

Application of Improved Satin Bowerbird Optimizer in Image Segmentation

*Linguo Li[1,2], Shunqiang Qian[1], Zhangfei Li[1] and Shujing Li[1]**

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

Aiming at the problems of low optimization accuracy and slow convergence speed of Satin Bowerbird Optimizer (SBO), an improved Satin Bowerbird Optimizer (ISBO) based on chaotic initialization and Cauchy mutation strategy is proposed. In order to improve the value of the proposed algorithm in engineering and practical applications, we apply it to the segmentation of medical and plant images. To improve the optimization accuracy, convergence speed and pertinence of the initial population, the population is initialized by introducing the Logistic chaotic map. To avoid the algorithm falling into local optimum (prematurity), the search performance of the algorithm is improved through Cauchy mutation strategy. Based on extensive visual and quantitative data analysis, this paper conducts a comparative analysis of the ISBO with the SBO, the fuzzy Gray Wolf Optimizer (FGWO), and the Fuzzy Coyote Optimization Algorithm (FCOA). The results show that the ISBO achieves better segmentation effects in both medical and plant disease images.

Keywords: satin bowerbird optimizer, chaotic initialization, Cauchy mutation strategy, medical image, plant image

OPEN ACCESS

Edited by:
Yu Xue,
Nanjing University of Information Science and Technology, China

Reviewed by:
Daqing Wu,
Shanghai Ocean University, China
Xijian Fan,
Nanjing Forestry University, China
Jianbo Meng,
Chongqing University, China

****Correspondence:***
Shujing Li
lishujing@fynu.edu.cn

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

Received: *08 April 2022*
Accepted: *22 April 2022*
Published: *06 May 2022*

Citation:
Li L, Qian S, Li Z and Li S (2022) Application of Improved Satin Bowerbird Optimizer in Image Segmentation. Front. Plant Sci. 13:915811. doi: 10.3389/fpls.2022.915811

INTRODUCTION

Image segmentation is an important means of image processing, analysis, understanding and computer vision. In addition to efficiently locating the region of interest in an image, it can also be widely used in the fields of feature extraction (Dobreva et al., 2021), image classification (Houssein et al., 2021; Xue et al., 2022), medical diagnosis (Kurmi et al., 2021) and plant disease segmentation (Akay et al., 2021). The quality of image segmentation is affected by many factors such as illumination change, inter-class difference and background complexity. There is still a considerable gap between existing image segmentation technology and the requirements of intelligent recognition and machine vision. Therefore, image segmentation remains a very open research field with unlimited potential. In recent times, a variety of image segmentation schemes have been proposed based on the needs of image understanding and machine vision, including thresholding, edge detection, and region methods (Pare et al., 2020; Mittal et al., 2021). The thresholding method divides the image into non-overlapping regions according to the gray level of the image through a threshold vector (one or more thresholds). Compared with other kinds of methods, such a method is easier to perform and involves less calculation. Therefore, it has been widely used in many fields, such as those involving medical images (Avola et al., 2021) and plant disease images (Neupane and Baysal Gurel, 2021).

Although the thresholding method is a simple and effective image segmentation procedure, the selection of optimal threshold remains crucial due to the influence of image background complexity, regional contrast and other factors. With the increase in the number of thresholds, the amount of calculation involved in this method also increases considerably. Therefore, the intelligent optimization algorithm based on specific objective function is widely used (Singh et al., 2021). Yazid et al. (2022) comprehensively analyze the effects of image brightness difference, target size and noise on Kapur and Li entropies in image thresholding. Through numerous experiments, the author verify the effectiveness of the two entropies, and put forward reasonable parameter setting suggestions for the application of the two entropies in image segmentation. Wu et al. (2021) take Kapur as the objective function and improve the preconditioning optimization algorithm (HPOA) through the evolutionary state strategy. In the experimental verification stage, the authors verify the excellent performance of Kapur entropy in color image segmentation through peak signal-to-noise ratio (PSNR), feature similarity index (FSIM) and structural similarity index (SSIM); this method is superior to other methods such as the moth flame optimization algorithm (MFO), and the multiverse optimizer (MVO). Sowjanya and Injeti (2021) verify the effects of Kapur and Otsu objective functions in multi-thresholding image segmentation through the butterfly optimization algorithm (BOA) and the gasses Brownian motion optimization (GBMO) fusion algorithm. Compared with seven optimization algorithms such as WOA (whale optimization algorithm) and SSO (social spider optimization), this method obtains better optimal threshold and further improves the image segmentation quality. Liu et al. (2021) use 2D Kapur entropy as the objective function and improve the differential evolution (DE) algorithm through slime mould foraging behavior, verifying the better performance of this method in breast cancer image segmentation. Experimental results show that this method not only improves convergence accuracy, but also reduces the risk of falling into local optimization. Kurban et al. (2021) use six intelligent optimization algorithms including equilibrium optimization (EO) and political optimizer (PO) to optimize Kapur and Otsu objective functions and obtain the optimal threshold. Through the evaluation indices such as PSNR, FSIM, Kapur-based marine predictors algorithm (MPA) and turbine flow of water-based optimization (TFWO) perform better than the other four methods in color aerial image segmentation. Anitha et al. (2021) take Kapur and Otsu as the objective function and obtain the optimal threshold through the modified white optimization algorithm (MWOA). Compared with GA, PSO and ABC algorithms, this method demonstrates superior convergence speed, feature stability and image segmentation quality. The time efficiency of this method is also higher. According to the specific objective function-based multi-level image thresholding methods in recent years, the use frequency of Kapur entropy is relatively high, and it also shows a reliable optimization effect in experimental comparison. Therefore, this paper continues to take the fuzzy Kapur (Li et al., 2016) as the objective function to analyze the performance of intelligent optimization algorithms in medical and plant image segmentation.

In view of the high computational complexity of multi-level thresholding and the demanding requirements of medical and plant image segmentation, the intelligent optimization algorithm is widely used to obtain the optimal threshold (Avola et al., 2021; Mittal et al., 2021; Neupane and Baysal Gurel, 2021). Li et al. (2016) improve the search strategy of gray wolf optimizer (GWO) by weighting the optimal population, enhancing the quality of image segmentation through median aggregation. The experimental verification of Berkeley segmentation dataset benchmarks 500 (BSD500) shows that this method obtains higher image segmentation accuracy in evaluation parameters such as PSNR and FSIM, compared with GWO, electromagnetism optimization (EO) and fuzzy DE. Su et al. (2022) improve the artificial bee colony algorithm (ABC) through horizontal and vertical search strategies. This algorithm improves the convergence speed and the quality of the optimal solution of ABC to some extent. In the experimental verification, the authors compare the performance of the algorithm with that of the original algorithm through 30 benchmark functions and apply the improved algorithm to Covid-19 X-ray images. In comparison with WOA, Sine Cosine Algorithm (SCA), Harris Hawks optimizer (HHO), Spherical Search optimizer (SSO) and other methods, the algorithm has better image segmentation effect on Covid-19 X-ray images with Kapur entropy as the objective function. Chen et al. (2021) improve the slim mount algorithm (SMA) through the threshold update mechanism of ABC and form a new ASMA fusion algorithm. Based on the analysis of 30 standard functions, this method obtains a better optimal solution and can effectively avoid prematurity. And the authors have achieved good results in the segmentation of standard images and lupus nephritis images. Li et al. (2021) inspired by the idea of differential evolution and make full use of the current number of iterations and the maximum number of iterations to improve the population search strategy of Coyote optimization algorithm (COA). Compared with COA, fuzzy ABC and fuzzy GWO, this method is better in visual and PSNR segmentation quality evaluations of BSD500 and medical image segmentation. Chakraborty and Mali (2022) improve the cuckoo search (CS) algorithm using fuzzy theory. Through the evaluation indices such as PSNR, SSIM, this method enjoys certain advantages in medical image segmentation quality and calculation time, compared with the improved DE, moth flame optimization algorithm and other methods. Singh et al. (2022) combine the tune swarm algorithm (TSA) and the naked mole rat algorithm (NMRA), and compare the data of CEC 2019 standard function and image segmentation, they show that this method is better than PSO, GA and others. According to the application of intelligent optimization algorithms in multi-level thresholding in recent years, we find this kind of method with numerous variations has been used widely in different fields. Taking account of the characteristics of medical and plant images, and the amount of experimental verification involved, this paper seeks to enhance the initialization and optimal population search strategy of the SBO. It further seeks to refine its optimization accuracy and convergence speed in image segmentation. Thus, image segmentation quality is improved.

The remainder of this paper is set out as follows: "Overview of the SBO" section briefly introduces the SBO. "The Improved

Satin Bowerbird Optimizer Based on Chaos Initialization and Cauchy Mutation" section presents the improved SBO (ISBO) based on chaotic initialization and Cauchy mutation strategy. "Selection of the Objective Function" section analyzes the selection of objective function corresponding to ISBO. "Comparison and Analysis of Experimental Results with the Medical and Plant Images" section details the performance of ISBO for medical and plant image segmentation. Finally, "Conclusion" section concludes this study.

OVERVIEW OF THE SBO

Satin Bowerbird Optimizer (SBO; Moosavi and Bardsiri, 2017) is an intelligent optimization algorithm that simulates the breeding behavior of an adult male Satin Bowerbird in the wild. As a wild bird with strong survival and reproduction skills, a mature male Satin Bowerbird wins the favor of the female by carefully constructing a courtship cabin, and attracts the opposite sex through continuous loud singing, holding a luminous object in its beak, to improve the probability of a successful courtship. In the courtship process the male should not only ensure a successful courtship cabin construction, but also constantly resist the challenges of its competitors, to prevent the nest from being damaged. According to Satin Bowerbird survival "rules," the SBO algorithm includes the following steps:

1. Random generation of the initial population of Satin Bowerbirds. An initial population of several *NB* individuals is randomly generated in a solution space. The position of each courtship cabin is defined as *D*-dimension, and the current population evolution algebra is *t*.
2. Calculate the fitness value (objective function) of each individual, and then calculate the ratio of the fitness value to the overall fitness value to represent the probability of the individual being selected. The probability of selecting the courtship cabin is calculated by Equation (1), $\mathrm{fit_i}$ represents the fitness value of the i-th courtship cabin, which can be calculated by Equation (2), $\mathrm{f}(\mathrm{x_i})$ represents the objective function value of the i-th courtship cabin.

$$Prob_i = \frac{fit_i}{\sum_{n=1}^{NB} fit_n} \tag{1}$$

$$fit_i = \begin{cases} \dfrac{1}{1+f(x_i)}, f(x_i) \geq 0 \\ 1+\left|f(x_i)\right|, f(x_i) < 0 \end{cases} \tag{2}$$

3. Update the population. According to the position information of the last iteration, the male Satin Bowerbird constantly adjusts the position of the courtship cabin, to continuously approach the optimal solution. The position update formula is as follows.

$$x_{ik}^{t+1} = x_{ik}^{t} + \lambda_k \left(\left(\frac{x_{jk} + x_{elite,k}}{2} \right) - x_{ik}^{k} \right) \tag{3}$$

Where $\mathrm{x_{ik}^t}$ is the k-dimensional component of the i-th individual in the t-th iteration; $\mathrm{x_{jk}}$ is the k-dimensional component of the optimal position ever found at present, $\mathrm{x_j}$ is determined by the roulette selection mechanism; $\mathrm{x_{elite,k}}$ is the k-dimensional component of the current global optimal position of the whole population. λ_k is the step factor, which is calculated by equation (4).

$$\lambda_k = \frac{\alpha}{1+P_j} \tag{4}$$

Where α is the maximum step size, and $\mathrm{P_j}$ is the probability of selecting the target courtship cabin, $\mathrm{P_j} \in [0,1]$. Equation (4) clearly shows that the greater the probability of selecting the target location, the smaller the step size. When the probability of selecting the target location is 0, the step size is the largest, denoted as α. When the probability of selecting the target location is 1, the step size is the smallest, denoted as $\alpha / 2$.

4. Individual variation to prevent it from falling into local optimization. A strong male often steals from other males' courtship cabins, and even destroys those cabins. Therefore, at the end of each iteration of the algorithm, there is a certain random mutation probability to improve the mutation of the algorithm. At this stage, $\mathrm{x_{ik}}$ the follows the normal distribution, as shown in Equation (5).

$$x_{ik}^{t+1} \sim N\left(x_{ik}^{t}, \sigma^2\right) \tag{5}$$

$$N\left(x_{ik}, \sigma^2\right) = x_{ik}^{t} + \left(\sigma * N(0,1)\right) \tag{6}$$

The calculation formula of the SD is as:

$$\sigma = z * \left(\mathrm{var_{max}} - \mathrm{var_{min}} \right) \tag{7}$$

Where z is the scaling factor, $\mathrm{var_{max}}$ and $\mathrm{var_{min}}$ are, respectively, the upper and lower limits of the variable $\mathrm{x_i}$.

5. At the end of each iteration, a new combination population is formed from the initial population and the population obtained from the mutation, and the fitness values of all individuals in the combination population are arranged in an ascending order. The individual with the largest objective function value is retained, and the other individuals are removed. If the end condition is satisfied at this stage, the optimal position and its corresponding optimal value will be outputted. Otherwise, the iteration will continue until the maximum number of iterations is reached.

THE IMPROVED SATIN BOWERBIRD OPTIMIZER BASED ON CHAOS INITIALIZATION AND CAUCHY MUTATION

Initialization of Logistic Chaos

Although according to the natural law, the initial population of the intelligent optimization algorithm adopts a random initialization mode, based on the purpose of the engineering application and convergence speed requirements, a better initialization method will greatly accelerate the convergence speed of intelligent optimization algorithm. The SBO also uses random values to initialize the population. Therefore, in this paper we introduce a Logistic chaotic map (Aniszewska, 2018) to improve the diversity of the initial population, thus obtaining a better initial population, and finally improving the optimization accuracy and convergence speed of the algorithm. The calculation method of Logistic chaotic map is shown in Equation (8):

$$X_{i+1} = \mu X_i .* \left(1 - X_i\right) \tag{8}$$

The value range of control parameters μ is set from 0 to 4. When the value of μ is larger, its chaos will be stronger. When the value of μ is 4, the chaotic initialization effect will be enhanced. Therefore, we take the value of μ at 4. Thus, the population initialization equation can be changed to Equation (9):

$$\mathrm{pop}(\mathrm{i}).\mathrm{Position} = \mathrm{Z}(\mathrm{i},:).*\left(\mathrm{VarMax} - \mathrm{VarMin}\right) + \mathrm{VarMin} \tag{9}$$

Where $\mathrm{Z}(\mathrm{i},:)$ represents $\mathrm{X_{i+1}}$ of Equation (8).

Cauchy Variation Strategy

The SBO is prone to fall into local optimization in its mutation stage. To solve this problem, this paper uses the Cauchy mutation strategy (Karakus et al., 2020) to replace the original mutation strategy of the SBO. The peak distribution of Cauchy function at the coordinate origin is shorter, but the distribution in the rest is longer. Using Cauchy mutation can produce greater disturbance nearby the current population. Compared with the original mutation strategy of the SBO, the improved method can produce greater and wider individual mutation, thus ensuring the flexibility and distinctiveness of mutation. The calculation equation of Cauchy variation strategy is shown in Equation (10):

$$X_{i,j}^{t+1} = X_{best}(t) + Cauchy(0,1) \oplus X_{best}(t) \tag{10}$$

Where *Cauchy* (0,1) represents the standard Cauchy distribution; $\mathrm{X_{best}}(\mathrm{t})$ is the position of an individual that needs variation.

The corresponding variation probability is calculated by Equation (11):

$$P_s = -\exp\left(1 - \frac{it}{MaxIt}\right)^{20} + p \tag{11}$$

Where the value of p is taken at 0.05, MaxIt represents the maximum number of iterations, and it represents the current number of iterations. If r and $<Ps$, the Cauchy mutation operation will not be performed. Otherwise, the mutation operation will continue.

SELECTION OF THE OBJECTIVE FUNCTION

According to the analysis of "Introduction"section, Kapur entropy is the commonly used objective function of multi-level thresholding. In Li et al. (2021), the Kapur entropy based on fuzzy theory is proposed, and the experiments show that the fuzzy Kapur entropy has a better optimization effect. Therefore, in order to reflect the optimal performance of the ISBO, the fuzzy Kapur continues to be set as the objective function in our work; next the optimal threshold is obtained through ISBO optimization, and the image is divided into multiple target regions. Assuming $\mathrm{I}(\mathrm{x,y})\mathrm{t}$ is the gray image that needs to be processed, $\mathrm{th_1, th_2, \ldots, th_d}$ are the preset D thresholds, then fuzzy Kapur represents the sum of probability statistics being divided into D+1 different gray distributions, which are shown in Equation (12):

$$H\left(th_1, th_2, \ldots, th_d\right) = H_0 + H_1 + \ldots + H_d \tag{12}$$

Where $\mathrm{H_i}$ is the entropy of the i-th gray distribution, which is expressed as:

$$\begin{aligned}
\boldsymbol{H_0} &= -\textstyle\sum_{i=0}^{L-1} \frac{\mu_0(i)*p_i}{\omega_0} \boldsymbol{ln}\left(\frac{\mu_0(i)*p_i}{\omega_0}\right) \\
\boldsymbol{\omega_0} &= \textstyle\sum_{i=0}^{L-1} \left(\boldsymbol{\mu_0(i)} * \boldsymbol{p_i}\right) \\
\boldsymbol{H_1} &= -\textstyle\sum_{i=0}^{L-1} \frac{\mu_1(i)*p_i}{\omega_1} \boldsymbol{ln}\left(\frac{\mu_1(i)*p_i}{\omega_1}\right) \\
\boldsymbol{\omega_1} &= \textstyle\sum_{i=0}^{L-1} \left(\boldsymbol{\mu_1(i)} * \boldsymbol{p_i}\right) \\
\boldsymbol{H_j} &= -\textstyle\sum_{i=0}^{L-1} \frac{\mu_j(i)*p_i}{\omega_j} \boldsymbol{ln}\left(\frac{\mu_j(i)*p_i}{\omega_j}\right) \\
\boldsymbol{\omega_j} &= \textstyle\sum_{i=0}^{L-1} \left(\boldsymbol{\mu_j(i)} * \boldsymbol{p_i}\right) \\
\boldsymbol{H_m} &= -\textstyle\sum_{i=0}^{L-1} \frac{\mu_m(i)*p_i}{\omega_m} \boldsymbol{ln}\left(\frac{\mu_m(i)*p_i}{\omega_m}\right) \\
\boldsymbol{\omega_m} &= \textstyle\sum_{i=0}^{L-1} \left(\boldsymbol{\mu_m(i)} * \boldsymbol{p_i}\right)
\end{aligned} \tag{13}$$

Where $\mu_0(\mathrm{i})$ is trapezoidal membership function (Li et al., 2016), and the threshold is calculated by fuzzy parameters, as shown in Equation (14):

$$th_1 = \frac{a_1 + a_2}{2}, th_2 = \frac{\mathrm{a_3} + a_4}{2}, \cdots, th_d = \frac{a_{n-1} + a_n}{2} \tag{14}$$

COMPARISON AND ANALYSIS OF EXPERIMENTAL RESULTS WITH THE MEDICAL AND PLANT IMAGES

Parameter Setting and Discussion

In order to verify the value of ISBO in engineering and practical applications, this paper focuses on the segmentation effect in

medical and plant images. Medical image segmentation can assist doctors to determine the location of lesions, especially in patients with mild disease, and can effectively help avoid doctors' diagnostic errors. Plant disease segmentation also has high application value in the field of agriculture. It can effectively detect pests and diseases and improve crop yield in a timely manner, especially in large-scale planting. This section first analyzes and compares the segmentation effects of medical images, and then gives the experimental results of plant images in the following section. The experiment of ISBO is performed on Windows10 (64bit), Intel Core i5 processor with 8GB RAM, using programming software MATLAB R2016a. The setting of the experimental parameters is shown in **Table 1**, like in Moosavi and Bardsiri (2017) and Li et al. (2021).

TABLE 1 | Experimental parameter setting.

Parameter	nPop	Alpha	Thresholds	Iterations
Value	20	0.94	2,3,5	1,000

TABLE 2 | Experimental results with different population numbers.

nPop	20	40	50	60	80
PSNR	26.1277	25.8615	25.4419	25.3696	24.9538
FSIM	0.9069	0.8951	0.8910	0.8890	0.8575

In order to prove that the parameter setting in **Table 1** can improve the optimization effect of ISBO, **Table 2** lists the PSNR and FSIM image segmentation quality evaluation values of Brn (**Figure 1A**) images with different population numbers, when the threshold is 5. According to the literature review in "Introduction" section, PSNR and FSIM are the two most commonly used evaluation indices in image segmentation. PSNR is called peak signal-to-noise ratio. The greater its value, the smaller the distortion of the image. FSIM is called feature similarity. The greater its value, the higher the similarity between the two images, and the better the quality of image segmentation. It can be observed from **Table 2** that when the population number is 20, the PSNR value and the FSIM value reach their maximum. With the increase in the population number, the values of PSNR and FSIM begin to decrease gradually. Therefore, in this paper, the population number of 20 is adopted.

In (Moosavi and Bardsiri, 2017), the step size is set to 0.94. In order to verify whether this step size is optimal, **Table 3** lists the impact of step threshold on image segmentation quality with the same image. **Table 3** demonstrates that when the value of the step threshold increases from 0.5 to 0.94, the PSNR value and FSIM value also increase. However, when

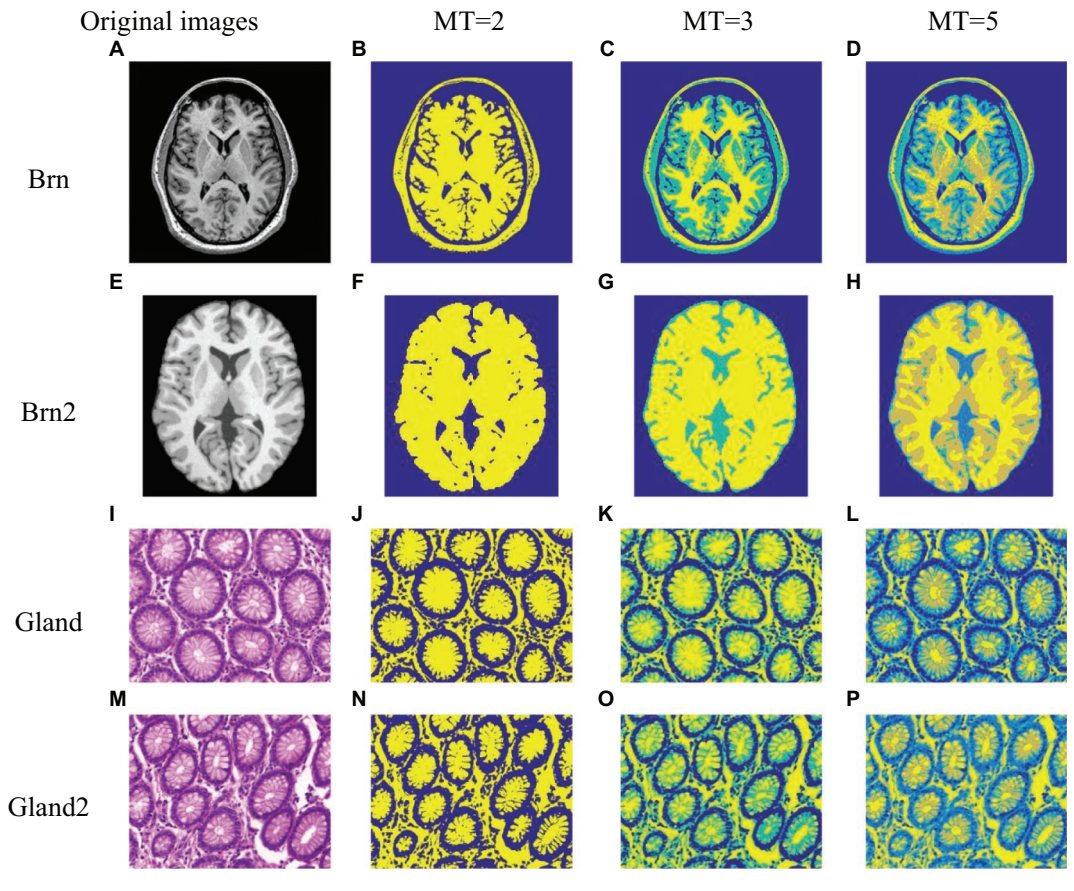


FIGURE 1 | The medical image segmentation results based on the ISBO. **(A-P)** only characterize the serial numbers of different experimental images.

the step threshold continues to increase from 0.94 onwards, the PSNR value and FSIM value begin to decrease. Then we can determine that the result is optimized, when the maximum step is 0.94. Therefore, this paper sets the threshold step to 0.94.

Table 4 lists the effects of the maximum number of iterations on image segmentation quality. When the maximum number of iterations is 1,000, the PSNR value and FSIM value of the algorithm in this paper are optimal, therefore we decide to set the maximum number of iterations of ISBO to 1,000.

Three spatial local information aggregation methods are cited in Li et al. (2016, 2021). Similarly, this paper also uses regional aggregation to improve the effect of image segmentation. To further compare the advantages and disadvantages of the proposed algorithm where the three aggregation methods are applied, **Figure 2** shows the segmentation results of Brn (**Figure 1A**) image by ISBO algorithm using the three aggregation methods, and **Table 5** lists the PSNR evaluation value obtained by using the proposed method in this paper under different thresholds.

Figures 2A–I demonstrates that the three aggregation methods have achieved satisfactory image segmentation results, but the visual contrast effect is not visibly obvious. In order to demonstrate the advantages and disadvantages of the three methods in detail, we can observe from **Table 5** that the median aggregation method obtains higher PSNR values under all thresholding conditions, compared with the other two methods. Therefore, ISBO is combined with the median aggregation method to complete the next experimental comparison and analysis.

TABLE 3 | Experimental results with different values of the step thresholds.

Alpha	0.5	0.94	1	1.5
PSNR	23.8823	26.1277	25.9621	24.8653
FSIM	0.8621	0.9069	0.8966	0.8593

TABLE 4 | Experimental results with different maximum iterations.

Iterations	500	1,000	3,000	5,000	10,000
PSNR	25.1937	26.1277	24.4850	25.1867	25.5288
FSIM	0.8604	0.9069	0.8547	0.8912	0.8791

Medical Image Experiment Results

Figures 1B–D, F–H, J–L and **N–P** show the image segmentation results of four medical images [Brain images: (A) and (E), Gland images: (I) and (M)] which collected from Simulated Brain Database (SBD)[1] and Gland Segmentation in Colon Histology Images (GlaS)[2] with different threshold numbers. In order to

[1]https://brainweb.bic.mni.mcgill.ca/brainweb/

[2]https://warwick.ac.uk/fac/cross_fac/tia

TABLE 5 | Experimental results with three different aggregation methods.

No. of thresholds	PSNR		
	Average	Iterative median	Median
2	15.6635	18.1473	18.2564
3	18.7338	22.5321	22.6985
5	25.8563	26.0101	26.2851

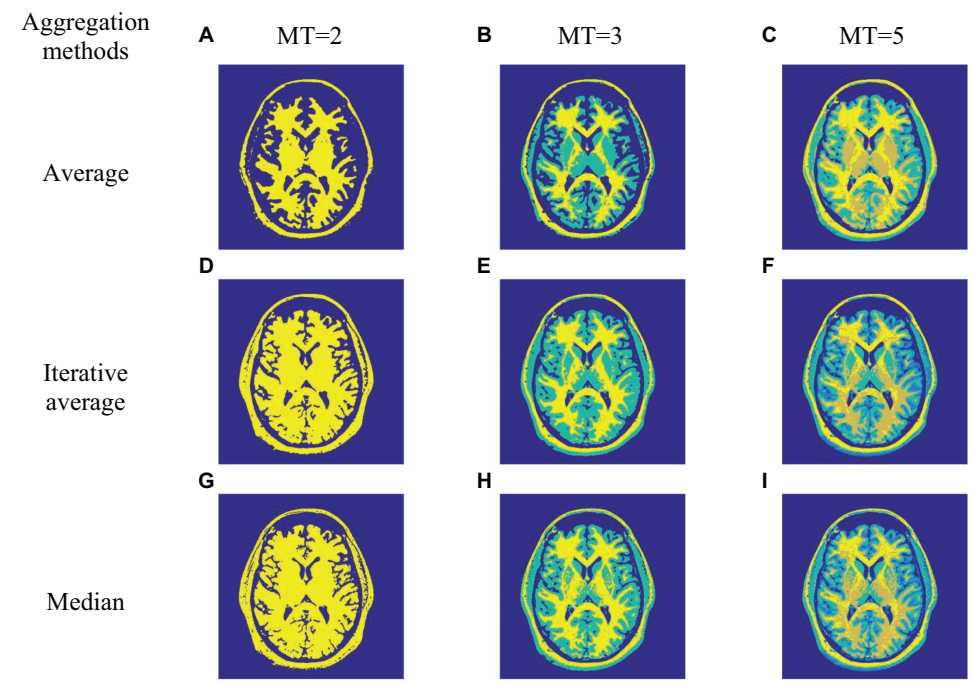


FIGURE 2 | Image segmentation effect of improved satin bowerbird optimizer (ISBO) based on three aggregation methods. **(A-I)** only characterize the serial numbers of different experimental images.

TABLE 6 | The experimental result data with medical images.

Image	MT	Thresholds					PSNR	FSIM
Brn	2	38.5	169.5				18.2564	0.6299
	3	25.5	117	196			22.6985	0.7868
	5	15.5	76	137	162	221.5	26.2851	0.9080
Brn2	2	71.5	175.5				17.6193	0.5881
	3	26.5	86	211.5			20.6821	0.6828
	5	21	69	104	160	219.5	26.2558	0.8210
Gland	2	111.5	175				13.3887	0.5657
	3	77.5	150.5	183			13.8454	0.7133
	5	59.5	135.5	156.5	195	236.5	14.1263	0.8328
Gland2	2	93.5	189.5				13.2126	0.5746
	3	70	142.5	205.5			13.8357	0.7349
	5	24	105.5	136.5	172	243.5	14.1201	0.8404

TABLE 7 | The experimental data with different algorithms.

Image	MT	PSNR			
		ISBO	SBO	FGWO	FCOA
Brn	2	18.2564	18.1473	18.1473	16.5122
	3	22.6985	22.5213	22.4268	21.4169
	5	26.2851	25.5405	25.5976	25.5721
Brn2	2	17.6193	17.5533	11.8691	13.5127
	3	20.6821	20.2781	19.5746	20.5012
	5	26.2558	25.1938	23.8653	25.2699
Gland	2	13.3887	13.2855	13.3207	13.3219
	3	13.8454	13.5825	13.7803	13.6153
	5	14.1263	14.1256	14.0953	14.0777
Gland2	2	13.2126	13.1912	13.2104	13.2140
	3	13.8357	13.6966	13.7586	13.7032
	5	14.1201	14.1110	14.1163	14.1113

more intuitively demonstrate the quality of image segmentation results, **Table 6** lists the threshold distribution, PSNR and FSIM of the four medical images with different threshold numbers.

As can be observed from **Figure 1**, for brain medical images and gland medical images, the ISBO algorithm can clearly segment the structures or tissues of different medical images, can effectively assist doctors in medical diagnosis, and can provide more refined pre-processing data for accurate diagnosis or disease prediction. In order to further compare it with other similar algorithms, as visual results alone cannot fully illustrate the problem, **Table 6** lists the threshold distribution and PSNR and FSIM evaluation data of the images in **Figure 1** after being optimized by the ISBO. From the data analysis, we can see that with the increase in the number of thresholds, the threshold distribution of the proposed method in our work tends to be balanced gradually, and better expected values are obtained in the PSNR and FSIM indices.

Quantitative Comparison and Analysis of Similar Algorithms

To fully illustrate the advantages of the proposed algorithm in this paper, based on **Figure 1**; **Tables 6, 7** lists the experimental data of ISBO and SBO, FGWO and FCOA, and analyzes the advantages and disadvantages of ISBO with PSNR as the evaluation standard.

The experimental data in **Table 7** demonstrates that compared with SBO, ISBO obtains higher PSNR value in all cases. Specifically, the PSNR value of ISBO algorithm increases by 0.2583 on average, − a mean percentage increase of 1.47%. The maximum increase stands at 1.062 and the minimum increase at 0.0007. Similarly, the maximum percentage increase is 6.03%, the minimum 0.004%. Compared to the FGWO, the PSNR value of the ISBO is also superior with an average increase of 0.8803 and an average increase ratio of 5.18%, of which the maximum increase is 5.7502 and the minimum 0.0022. By percentage, the maximum increase is 33.86% and the minimum 0.01%. Finally, compared with the FCOA, the ISBO is better than the FCOA except when the Gland 2 image threshold is 2. To be exact, the PSNR value of the ISBO is 0.8635 higher than that of the FCOA, with an average increase ratio of 4.96%. The highest increase occurs when the threshold of Brn2 image is 2, with an increase of 4.1066, and an increase ratio of 23.57%.

Comparison and Analysis of Plant Image Segmentation Experiments

In order to verify the superiority of ISBO and fully reflect its practical application value, we apply this algorithm to the plant image segmentation. In our study, four plant images in the Kaggle plant image dataset[3] are selected for experiment, and the parameter setting is consistent with that discussed in "Parameter Setting and Discussion" section. **Figures 3B–D, F–H, J–L** and **N–P** show the image segmentation results of the four plant images (**Figures 3A,E,I,M**) by ISBO with different threshold numbers. The corresponding **Table 8** lists the experimental result data.

Figure 3 shows that ISBO achieves more refined plant image segmentation, and effectively realizes the segmentation of lesion area. Like medical image analysis, it cannot effectively make the comparative analysis with other similar algorithms in a visual way. Therefore, **Table 8** lists the threshold distribution and PSNR and FSIM quantitative data corresponding to the plant image segmentation. **Table 8** shows that with the increase in the number of thresholds, the threshold vectors are more evenly distributed; the values of PSNR and FSIM are higher,

[3]https://www.kaggle.com/datasets/asheniranga/leaf-disease-dataset-combination

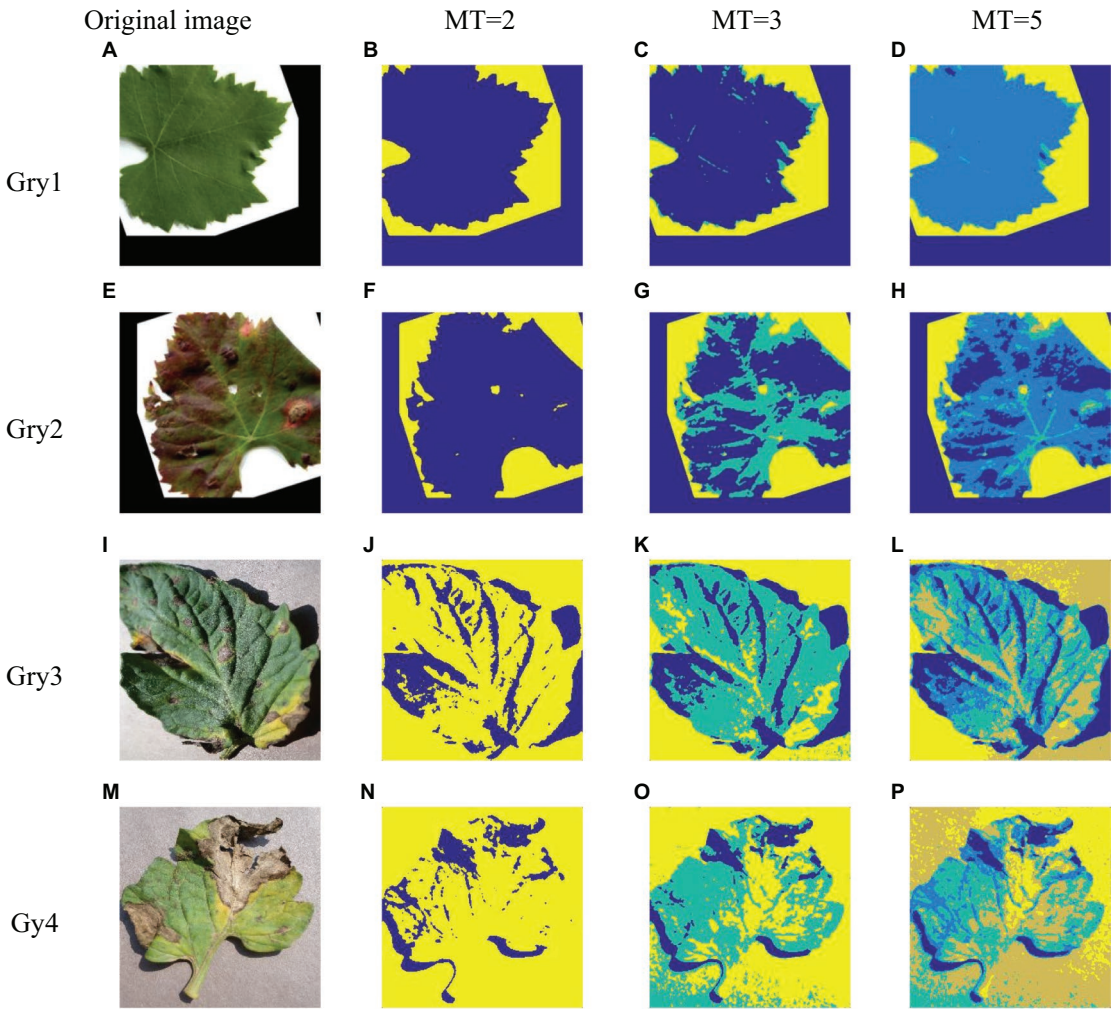


FIGURE 3 | The plant image segmentation result based on ISBO.

TABLE 8 | Experimental data of the plant image segmentation.

Image	MT	Thresholds					PSNR	FSIM
Gry1	2	105	187				16.2897	0.7236
	3	82.5	123.5	183.5			16.6751	0.7703
	5	27.5	76	144.5	178.5	226.5	26.1164	0.8181
Gry2	2	112	197				16.5825	0.6447
	3	51	101	175			19.0693	0.6779
	5	36	83.5	125	182	225.5	22.1451	0.7327
Gry3	2	43.5	135.5				16.4277	0.5099
	3	37	117	194.5			19.2348	0.5989
	5	37	92.5	125	165	232.5	20.7436	0.7151
Gry4	2	66.5	180.5				20.1022	0.5137
	3	59	138.5	185.5			22.0812	0.6039
	5	23.5	106.5	150	171	200.5	23.5093	0.7078

thus proving that the proposed algorithm in our work can effectively realize the segmentation of plant disease images.

To more intuitively evaluate the advantages and disadvantages of ISBO in plant image segmentation, this section also compares ISBO with SBO, FGWO and FCOA. **Table 9** lists the experimental result data of the four algorithms. It shows that the PSNR value of ISBO is superior to that of SBO in all cases. Specifically, compared with SBO, the PSNR value of the ISBO increases

TABLE 9 | Comparison of experimental data of the plant images.

Image	MT	PSNR			
		ISBO	SBO	FGWO	FCOA
Gry1	2	16.2897	16.2725	16.2854	16.2804
	3	16.6751	16.3979	16.5253	16.5139
	5	26.1164	26.0164	16.8329	16.6779
Gry2	2	16.5825	16.5326	16.5798	16.5732
	3	19.0693	16.9781	17.4758	17.1470
	5	22.1451	18.5860	18.7410	18.6002
Gry3	2	16.4277	16.3589	16.4068	16.0568
	3	19.2348	18.9205	18.7574	17.9570
	5	20.7436	20.5342	20.8043	20.6044
Gry4	2	20.1022	19.8106	20.0157	19.9443
	3	22.0812	21.4377	21.9520	21.8896
	5	23.5093	23.0684	23.1221	22.5566

by 0.6719 on average, and the average increase percentage is 3.49%. And the maximum increase reaches 3.5591 and the minimum increase reaches 0.0172. By percentage, the maximum increase reaches 18.5%, the minimum 0.09%. Compared with the FGWO, the ISBO is superior under other conditions except when the threshold of Gry3 is 5. Specifically, the PSNR value of the ISBO is 1.4126 higher than that of the FGWO, and the average increase percentage is 7.67%. In the case of the Gry3 image when the exceptional threshold 5 is set, the PSNR value of the ISBO is only 0.0607 lower than that of the FGWO. Finally, compared with the FCOA, the ISBO still obtains higher PSNR value in all cases. Similarly, compared with the FCOA, the PSNR value of the ISBO increases by 1.5146 on average, with an average increase percentage of 8.23%, of which the highest and lowest increase percentages are 51.3% and 0.05%, respectively.

CONCLUSION

In order to fully verify the value of the intelligent optimization algorithm based on specific objective function in medical and plant image segmentation applications, this paper introduces the ISBO into the multi-level thresholding of medical and plant images. The algorithm takes fuzzy Kapur as the objective function and optimizes a set of thresholds by improving the SBO algorithm to complete the initial image segmentation. On this basis, the median aggregation method is introduced to avoid the problem of over-segmentation or segmentation of outliers. In the process of improving the SBO algorithm, chaos initialization and Cauchy mutation strategy are also introduced to improve the convergence speed of the algorithm and reduce the risk of falling into local optimization. To prove the superiority of ISBO, this paper compares ISBO with SBO, FGWO and FCOA using medical and plant images. Through the comparison of visual and quantitative data, it can be observed that ISBO is more effective in the segmentation of medical and plant images.

DATA AVAILABILITY STATEMENT

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

AUTHOR CONTRIBUTIONS

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

FUNDING

This paper is supported by the National Youth Natural Science Foundation of China (61802208), the National Natural Science Foundation of China (61572261), the Natural Science Foundation of Anhui (1908085MF207, KJ2020A1215, KJ2021A1251, and KJ2021A1253), the Excellent Youth Talent Support Foundation of Anhui (gxyqZD2019097 and gxyqZD2021142), the Postdoctoral Foundation of Jiangsu (2018K009B), and the Foundation of Fuyang Normal University (TDJC2021008).

ACKNOWLEDGMENTS

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

REFERENCES

Akay, R., Saleh, R. A. A., Farea, S. M. O., and Kanaan, M. (2021). Multilevel thresholding segmentation of color plant disease images using metaheuristic optimization algorithms. *Neural Comput. Applic.* 34, 1161–1179. doi: 10.1007/s00521-021-06437-1

Aniszewska, D. (2018). New discrete chaotic multiplicative maps based on the logistic map. *Int. J. Bifurc. Chaos* 28:1850118. doi: 10.1142/S0218127418501183

Anitha, J., Pandian, S. I. A., and Agnes, S. A. (2021). An efficient multilevel color image thresholding based on modified whale optimization algorithme. *Expert Syst. Appl.* 178:115003. doi: 10.1016/j.eswa.2021.115003

Avola, D., Cinque, L., Fagioli, A., Foresti, G., and Mecca, A. (2021). Ltrasound medical imaging techniques: a survey. *ACM Comput. Surv.* 54, 1–38. doi: 10.1145/3447243

Chakraborty, S., and Mali, K. (2022). Biomedical image segmentation using fuzzy multilevel soft thresholding system coupled modified cuckoo search. *Biomed Signal Process.* 72:103324. doi: 10.1016/j.bspc.2021.103324

Chen, X., Huang, H., Heidari, A. A., Sun, C., Lv, Y., and Gui, W. (2021). An efficient multilevel thresholding image segmentation method based on the slime mould optimizer with the bee foraging mechanism: a real case with lupus nephritis images. *Comput. Biol. Med.* 142:105179. doi: 10.1016/j.compbiomed.2021.105179

Dobreva, I. D., Ruiz Guzman, H. A., Barrios Perez, I., Adams, T., Teare, B. L., Payton, P., et al. (2021). Thresholding analysis and feature extraction from

3D ground penetrating radar data for noninvasive assessment of peanut yield. *Remote Sens.* 13, 1–26. doi: 10.3390/rs13101896

Houssein, E. H., Emam, M. M., and Ali, A. A. (2021). An efficient multilevel thresholding segmentation method for thermography breast cancer imaging based on improved chimp optimization algorithm. *Expert Syst. Appl.* 185:115651. doi: 10.1016/j.eswa.2021.115651

Karakus, O., Mayo, P., and Achim, A. (2020). Convergence guarantees for non-convex optimisation with Cauchy-based penalties. *IEEE Trans. Signal Process.* 68, 6159–6170. doi: 10.1109/TSP.2020.3032231

Kurban, R., Durmus, A., and Karakose, E. (2021). A comparison of novel metaheuristic algorithms on color aerial image multilevel thresholding. *Eng. Appl. Artif. Intell.* 105:104410. doi: 10.1016/j.engappai.2021.104410

Kurmi, Y., Chaurasia, V., and Kapoor, N. (2021). Histopathology image segmentation and classification for cancer revelation. *Signal Image Video Process.* 15, 1341–1349. doi: 10.1007/s11760-021-01865-x

Li, L., Sun, L., Kang, W., Guo, J., Han, C., and Li, S. (2016). Fuzzy multilevel image thresholding based on modified discrete grey wolf optimizer and local information aggregation. *IEEE Access* 4, 6438–6450. doi: 10.1109/ACCESS.2016.2613940

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

Liu, L., Zhao, D., Yu, F., Heidari, A. A., Ru, J., Chen, H., et al. (2021). Performance optimization of differential evolution with slime mould algorithm for multilevel breast cancer image segmentation. *Comput. Biol. Med.* 138:104910. doi: 10.1016/j.compbiomed.2021.104910

Mittal, H., Pandey, A. C., Saraswat, M., Kumar, S., Pal, R., and Modwel, G. (2021). A comprehensive survey of image segmentation: clustering methods, performance parameters, and benchmark datasets. *Multimed. Tools Appl.* 1174, 1–26. doi: 10.1007/s11042-021-10594-9

Moosavi, S. H. S., and Bardsiri, V. K. (2017). Satin bowerbird optimizer: a new optimization algorithm to optimize ANFIS for software development effort estimation. *Eng. Appl. Artif. Intell.* 60, 1–15. doi: 10.1016/j.engappai.2017.01.006

Neupane, K., and Baysal Gurel, F. (2021). Automatic identification and monitoring of plant diseases using unmanned aerial vehicles: a review. *Remote Sens.* 13:3841. doi: 10.3390/rs13193841

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

Singh, S., Mittal, N., Singh, U., Salgotra, R., Zaguia, A., and Singh, D. (2022). A novel hybrid tunicate swarm naked mole-rat algorithm for image degmentation and numerical optimization. *Comput. Mater. Contin.* 71, 3445–3462. doi: 10.32604/cmc.2022.023004

Singh, S., Mittal, N., Thakur, D., Singh, H., Oliva, D., and Demin, A. (2021). Nature and biologically inspired image segmentation techniques. *Arch. Comput. Methods Eng.* 29, 1415–1442. doi: 10.1007/s11831-021-09619-1

Sowjanya, K., and Injeti, S. K. (2021). Investigation of butterfly optimization and gases brownian motion optimization algorithms for optimal multilevel image thresholding. *Expert Syst. Appl.* 182:115286. doi: 10.1016/j.eswa.2021.115286

Su, H., Zhao, D., Yu, F., Heidari, A. A., Zhang, Y., and Chen, H. (2022). Horizontal and vertical search artificial bee colony for image segmentation of COVID-19 X-ray images. *Comput. Biol. Med.* 142:105181. doi: 10.1016/j.compbiomed.2021.105181

Wu, B., Zhu, L., Cao, J., and Wang, J. (2021). A hybrid preaching optimization algorithm based on Kapur entropy for multilevel thresholding color image segmentation. *Entropy* 23:1599. doi: 10.3390/e23121599

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

Yazid, H., Basah, S. N., Rahim, S. A., Safar, M. J. A., and Basaruddin, K. S. (2022). Performance analysis of entropy thresholding for successful image segmentation. *Multimed. Tools Appl.* 81, 6433–6450. doi: 10.1007/s11042-021-11813-z

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: 19 May 2022
doi: 10.3389/fpls.2022.844522

A Segmentation-Guided Deep Learning Framework for Leaf Counting

Xijian Fan[1], Rui Zhou[1], Tardi Tjahjadi[2], Sruti Das Choudhury[3] and Qiaolin Ye[1]*

[1] College of Information Science and Technology, Nanjing Forestry University, Nanjing, China, [2] School of Engineering, University of Warwick, Coventry, United Kingdom, [3] Department of Biological Systems Engineering, University of Nebraska-Lincoln, Lincoln, NE, United States

Deep learning-based methods have recently provided a means to rapidly and effectively extract various plant traits due to their powerful ability to depict a plant image across a variety of species and growth conditions. In this study, we focus on dealing with two fundamental tasks in plant phenotyping, i.e., plant segmentation and leaf counting, and propose a two-steam deep learning framework for segmenting plants and counting leaves with various size and shape from two-dimensional plant images. In the first stream, a multi-scale segmentation model using spatial pyramid is developed to extract leaves with different size and shape, where the fine-grained details of leaves are captured using deep feature extractor. In the second stream, a regression counting model is proposed to estimate the number of leaves without any pre-detection, where an auxiliary binary mask from segmentation stream is introduced to enhance the counting performance by effectively alleviating the influence of complex background. Extensive pot experiments are conducted CVPPP 2017 Leaf Counting Challenge dataset, which contains images of Arabidopsis and tobacco plants. The experimental results demonstrate that the proposed framework achieves a promising performance both in plant segmentation and leaf counting, providing a reference for the automatic analysis of plant phenotypes.

Keywords: plant phenotyping, segmentation, deep CNN architecture, leaf counting, multiple traits

OPEN ACCESS

Edited by:
Yu Xue,
Nanjing University of Information Science and Technology, China

Reviewed by:
Marcin Wozniak,
Silesian University of Technology, Poland
Li Chaorong,
Yibin University, China

****Correspondence:***
Xijian Fan
xijian.fan@njfu.edu.cn

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

Received: *11 January 2022*
Accepted: *19 April 2022*
Published: *19 May 2022*

Citation:
Fan X, Zhou R, Tjahjadi T, Das Choudhury S and Ye Q (2022) A Segmentation-Guided Deep Learning Framework for Leaf Counting. Front. Plant Sci. 13:844522. doi: 10.3389/fpls.2022.844522

INTRODUCTION

Plant phenotype is a set of observable traits of a plant, which is heavily influenced by the interaction between plant gene expression and environmental factor (Siebner et al., 2009). The accurate and efficient monitoring of phenotypes is essential for plant cultivation, which is a prerequisite for intelligent production and planting, and information/data management. The traditional monitoring of plant phenotype mainly requires manual observation and measurement to analyse the appearance of plants in terms of their shape, texture, colour, and other characteristic morphological phenotypes (Montero et al., 2000; Minervini et al., 2015). Such an approach is labour intensive, which is time-consuming and prone to error due to the reliance on subjective perception (Yang et al., 2020). Image-based plant phenotyping allows non-invasive and distant observation, reducing the effects of manual interference and vastly increasing the scale and throughput of plant phenotyping activities. However, it still requires a robust algorithm to automatically process the

input image to provide accurate and reliable phenotypic estimation (Scharr et al., 2016). In addition, such an algorithm should be able to estimate a wide diversity of phenotypes, which allows for a range of different scientific applications. The current trend of image-based plant phenotyping attempts to combine image processing (e.g., noise removal and image enhancement), feature extraction and machine learning to obtain effective and efficient estimation (Tsaftaris et al., 2016). In recent years, deep learning-based methods have made remarkable progress in the field of computer vision such as semantic segmentation, classification, and object detection (Lecun et al., 2015). They integrate feature extraction and classification using a single convolutional neural network (CNN) based framework, which is trained in an end-to-end fashion. Due to their powerful ability to capture meaningful feature representation, deep learning-based methods are drawing more attention in the plant research community (Dhaka et al., 2021; Kundu et al., 2021) and have also been applied to deal with different tasks in plant phenotyping (Choudhury et al., 2019).

Plant segmentation and leaf counting are two fundamental tasks of plant phenotyping as they are relevant to the developmental stage of a plant, and are considered essential means of providing vital indicators for the evaluation of plant growth (e.g., growth regulation and flowering time), yield potential, and plant health. Moreover, they help farmers and horticulturists to make better decision regarding cultivation strategic and timely horticulture adjustments. Plant segmentation aims to extract the plant area, shape, and size from a visual perspective by segmenting an entire plant from the scene background in an image. Such a task closely relates to the semantic/instance segmentation problems, and some researchers have addressed this task using instance/semantic segmentation (Romera-Paredes and Torr, 2016; Ren and Zemel, 2017; Ward et al., 2018; Zhu et al., 2018), achieving promising performance. Leaf counting aims to estimate the precise number of leaves of a plant. There are two mainstream ways to infer the leaf count or leaf number: (1) estimating the leaf number as a sub-product of leaf segmentation or detection (Girshick, 2015; Kong et al., 2020; Kumar and Domnic, 2020; Lin and Guo, 2020; Lu and Cao, 2020; Tassis et al., 2021); and (2) directly regarding the task as a holistic regression problem (Dobrescu et al., 2017; Giuffrida et al., 2018; Itzhaky et al., 2018; Ubbens et al., 2018; Mishra et al., 2021). The methods have successfully addressed the tasks of leaf segmentation and counting using machine learning and especially deep learning methods, which uncover the intrinsic information from plant images, even when they contain complex structure. However, they merely focus on a single task, i.e., learn one plant trait at a time. Thus, they might ignore the facts that plant phenotype traits tend to be associated with each other and lack the insight to the potential relationship between different traits (Gomes and Zheng, 2020). For instance, the leaf number is associated with the leaf area, age, and genotype. We believe that incorporating multiple traits in the deep CNN architecture could be beneficial for learning more reliable and discriminative information than using only one trait. Dobrescu et al. (2020) presented a multi-task framework for leaf count, projected leaf area, and genotyping, where they compute three plant traits at the same time by using the share representation layers. However, they did not address the tasks of plant segmentation that is more challenging due to the requirement of classifying all the leaves (foreground) pixel by pixel.

Convolutional neural network based methods have been applied to plant and leaf segmentation in plant phenotyping. Aich and Stavness (2017) used a CNN based deconvolutional network for plant (foreground) and leaf segmentation. Kuznichov et al. (2019) utilised data augmentation technology to maintain the geometric structure and physical appearance of plant in images to improve the leaf segmentation. Bell and Dee (2019) employed a relatively shallow CNN model to classify image edges extracted using Canny edge detector, which distinguished the occluding pairs of leaves. Ren and Zemel (2017) adopted recurrent neural network (RNN) to generate a single segmented template for each leaf and combined convolutional long short-term memory (LSTM) network using spatial inhibition modules. They then used dynamical non-maximal suppression to leverage the previously segmented instances to enhance the segmentation. Although achieving promising results, these methods use the shallow CNN model, which is inadequate to capture the meaningful information of the diversity of plant images. Moreover, all methods concentrate on addressing the single task, i.e., leaf/plant segmentation in an independent pipeline.

Image segmentation using deep learning has gained a significant advance, and a few benchmark methods have been proposed. Fully convolutional networks (FCN) (Long et al., 2015) and U-Net (Ronneberger et al., 2015) are two representative models that are based on the encoder-decoder network architecture. Both of them share a similar idea, i.e., using skip connection, that shows the capability to capture the fine-grained characteristics of the target images. FCN summed the up-sampled feature maps with feature maps skipped from the encoder, while U-Net modified the way of feature concatenation by adding convolutions and non-linearities during each up-sampling step. Another mainstream work is using spatial pyramid pooling ideas. PSPNet employed a pyramid parsing operation that captures global context information by region feature aggregation (Zhao et al., 2017). DeepLab (Chen et al., 2017) introduced the atrous convolution with up-sampling filter for feature extraction, and extended it using spatial pyramid pooling to encode the multi-scale contextual semantics. However, the various scale pooling operations tend to lose local spatial details and will fail to maintain leaf target with high density if a small input size is adopted. The Mask Region Convolutional Neural Network (Mask-RCNN), proposed by He et al. (2017), extended the region proposal network by integrating a branch to predict segmentation mask on each ROI. Mask RCNN can segment the object with pixel-wise mask from a complicated background, which is suitable for the leaf segmentation. Thus, we developed our network model based on the backbone architecture in Mask-RCNN and simply replaced the plain skip connection with a nested dense skip pathway to enhance the ability to extract more fine-grained features in leaf images.

Leaf counting is also an important task in plant phenotyping, since leaf count is considered as an indicator for yield potential and plant health (Rahnemoonfar and Sheppard, 2017). From the

perspective of computer vision, leaf counting can be addressed along two different lines: (1) Regarding leaf counting as the sub-product of leaf segmentation or detection, leading to the leaf number following the segmentation module; and (2) Directly learning an image-to-count model to estimate the leaf number using training samples.

Direct Count

Leaf counting is regarded as a holistic regression task, in which a counting model estimates the leaf number for a given plant image. In this way, the machine learning based regression model solely needs the annotation of leaf number, which is an easier way to obtain compared with the pixel-wise annotations using segmentation. Dobrescu et al. (2017) presented a counting framework employing the ResNet50 backbone (He et al., 2016), in which the learning of leaf counting is performed by gathering samples from multiple sources. Itzhaky et al. (2018) proposed to estimate the leaf number using multi-scale representations and fuse them to make the final predictions. Ubbens et al. (2018) presented an open-source platform which aims to introduce a more generalised system for plant breeders, which can be used to count leaves across different datasets, as well as to assist other tasks e.g., projected leaf area and genotype classification. da Silva and Goncalves (2019) constructed a CNN based regression model to learn from images, where the skip connections of Resent50 (He et al., 2016) are considered efficient for leaf counting. Direct count could be a natural and easy selection as it is not necessary to annotate the image when training.

Counting via Detection or Segmentation

This approach regards the leaf counting problem as a sub-product of detection or segmentation, where the exact locations and number of the leaves are also obtained after detection or segmentation. Romera-Paredes and Torr (2016) proposed to learn an end-to-end segmentation model using RNN, that segments each leaf sequentially and then estimate the number of segmented leaves. Aich and Stavness (2017) used a CNN based deconvolutional network for leaf segmentation and a convolutional network for leaf counting. Kumar and Domnic (2019) developed a counting model with the combination of CNN and traditional methods, where graph-based method is used for U-Net segmentation and CNN-based is then used for leaf counting via a fine-tuned AlexNet. Ren and Zemel (2017), propose a neural network using which visual attention operation to jointly learn the instance segmentation and counting model, where sequential attention using LSTM cell is created by using temporal chain to output one instance at a time. However, such a segmentation or detection-based method has one limitation for counting. That is, only successfully segmented leaves are counted, and imperfect detection will result in reduced accuracy in counting. Unlike the aforementioned methods, we employ the segmented binary image to guide the learning of leaf counting, i.e., not counting directly from the segmented image, thus avoiding the effect of inaccurate detection or segmentation on the counting task.

In this study, we present in this article a two-stream framework, one stream for plant segmentation and the other stream for leaf counting based on regression. The resultant mask from segmentation stream is leveraged to guide the learning of leaf counting, which help to alleviate the inference of complex background. In order to obtain more semantic and meaningful feature representation of plant images, we employ the deep CNN as the model backbones of both two streams. By using the CNN paradigm, the two-stream model is robust and generalizes well regardless of the plant species and the quality of the acquired image data. This is achieved by one stream task supervising the training of the other stream task *via* sharing certain knowledge. To this end, we employ the segmented binary mask from the plant segmentation stream as an auxiliary cue to optimise the training process of the leaf counting stream. Introducing the binary mask to supervise the learning of leaf counting is based on two issues that exclusively exist in plant leaf counting: (1) some leaves might be partially occluded by other leaves, or are incomplete and fragmentary on their own, making them difficult to detect; and (2) the leaves sometimes contain complex background, increasing the challenge in leaf counting. These two issues led to incorrect or missing count where the meaningful and useful information of leaf is hard to maintain during the leaf counting. The binary mask effectively deals with these two issues by precisely locating all individual leaves while alleviating the effect of complex background. In addition, the binary mask of image samples brings more diversity of the input images by increasing the number of samples, which could be regarded as an implicit data augmentation.

Specifically, in our proposed framework, a two-stream deep neural network model segments the leaves and counts the number of leaves, where the segmented binary mask is employed as an auxiliary cue to supervise the learning of leaf counting. In the stream for segmentation, a multi-scale-based segmentation network is proposed to extract fine-grained characteristics of leaves. In the stream for leaf counting, we propose to learn a regression model based on the fine-tuned CNN model. During the learning of leaf counting, the segmented mask is utilized to highlight the target leaf region (foreground) of interest (ROI) from the entire image by removing the disturbance of complex background (i.e., non-leaf area, thus facilitating the counting process.

The contributions of this study are summarized as follows:

1. We propose to explore fine-grained characteristics, i.e., high inter-class similarity and low intra-class variations, widely existing in high throughput plant phenotyping that cause the failure in localizing the leaves within a small area during segmentation. To address this issue, we introduce a multi-scale U-Net segmentation model which compensates the upper-lower semantics difference by concatenating features in various scales. This model is learned in an end-to-end fashion, allowing for efficient segmentation of the leaves with different areas.
2. We propose a two-stream network based on deep CNN architecture to complete the leaf counting together with plant segmentation, in which the model outputs the segmentation results and directly estimates the leaf number.

3. We enhance the leaf counting by introducing the auxiliary binary information. The binary mask is utilised to supervise the leaf counting, which increases the contrast between the leaf target from background interference, and significantly aids the convergence of the counting regression model.

The remainder of the article is presented as follows: we review related work in Section "Introduction," present our method in Section "Proposed Method," provide the experimental results in Section "Experiments" and discuss the conclusions and further work in Section "Conclusion."

PROPOSED METHOD

We present a parallel two-stream network for determining leaf count and undertake segmentation simultaneously for the rosette-shaped plants as shown in **Figure 1**. The stream for segmentation adopts the nested U-Net (U-Net++) architecture (Zhou et al., 2018) as backbone to extract the target leaf region from the entire image using a binary mask. The stream for leaf counting learns the CNN based regression model which is customized by modifying its last layer to directly estimate the number of leaves where the segmented mask and original colour images with the leaf number label are mixed as input of the regression model. The streams for plant segmentation and count are designed separately first. The segmented binary mask denoting the area of the leaf is used as a complementary cue to supervise the learning of the count regression stream. This is because the two key traits of the two streams, i.e., the area and leaf number are often related to each other. Incorporating the leaf area into the estimation of leaf number during the learning of deep neural network aids not only to learn more meaningful and essential information, but also alleviates the influence of complex background.

Plant Segmentation Module

The segmentation module aims to extract the whole leaf area from the background. In order to enhance the robustness and accuracy of extraction, it is a necessity for the module to be in capacity to depict the characteristics existing in a plant image, i.e., fine-grained and variation in shape and size. To this end, we consider the nested U-Net as our backbone network for the segmentation. The nested U-Net model is proposed based on the U-Net that was originally proposed to meet the requirement on accurately segmenting medical images. Compared with the original U-Net model proposed by Ronneberger et al. (2015), the nested U-Net architecture replaces the plain skip connection with nested and dense skip connections, which can capture fine-grained information of the object in an image. Moreover, due to the up-sampling scheme, the U-Net model could locate leaves with different size and shape by using feature maps with different scales. By dealing with the characteristics in leaves, the nested U-Net is thus suitable for plant segmentation. Another problem needs to be addressed during training, namely the ROIs of plant segmentation comprise a relatively small segment of the entire image. Thus, negative samples (i.e., background pixels) are much larger than positive samples (i.e., leaf pixels), which resulted in an unbalanced binary classification problem. To address the problem, we integrate the binary cross-entropy (BCE) loss with dice loss together, and jointly guide the learning process of the segmentation. Generally, the nested U-Net consists of three main modules: encoding, decoding, and cross-layers dense concatenation. The feature maps in the same size are defined to be of the same layer, denoting the layers as L1–L5 from top to bottom. Each node represents a feature extraction module consisting of two 3×3 convolutional layers, followed by a rectified linear unit (ReLU) and a 2×2 max pooling that use stride 2 for down-sampling.

The output features from encoder are fused with the next encoder layer *via* up-sampling features across layers from top to bottom. The fusion outputs are concatenated with the corresponding up-sampled features of the next layer, and the process is iterated until there is no corresponding module in the next layer. The integrated feature maps are defined as

$$x^{i,j} = \begin{cases} \mathcal{H}\left(x^{i-1,j}\right) & j = 0 \\ \mathcal{H}\left(\left[\left[x^{i,k}\right]_{k=0}^{j-1}, \mathcal{U}\left(x^{i+1,j-1}\right)\right]\right) & j > 0 \end{cases} \quad (1)$$

where $\mathcal{H}(\cdot)$ denotes a convolution operation followed by an activation function, $\mathcal{U}(\cdot)$ denotes an up-sampling layer, and [] denotes the concatenation layer. Nodes at level $j = 0$ only receive input from the previous encoder layer; nodes at level $j = 1$ receive the encoder and sub-network input from two consecutive levels; and nodes $j > 1$ receive $j + 1$ inputs of which j inputs are the outputs of the previous j nodes in the same skip pathway and the last input is the up-sampled output from the lower skip pathway.

The dense skip connections between layers in the same dimension pass the output of the current module to all subsequent modules and fuse it with other input features. Thus, the overall U-Net++ feature fusion structure is in the form of an inverted pyramid, where the intermediate layer contains more accurate localisation information, while the in-depth layer captures pixel-level category information.

As a typical binary classification task, the core objective is to segment the plant image into a binary image by labelling the foreground and background pixels as 1 and 0, respectively. To overcome the class imbalance problem, BCE loss and Dice loss are combined to form the objective function to optimize the imbalance between the foreground and background pixels through back-propagation. Dice coefficient is a measure of the pixel degree of an ensemble, and the original expression takes the form of

$$d = \frac{2|X \cap Y|}{|X| + |Y|} \quad (2)$$

where X and Y are sets, and $s \in [0, 1]$, and the size of s reflects the similarity between the sets X and Y.

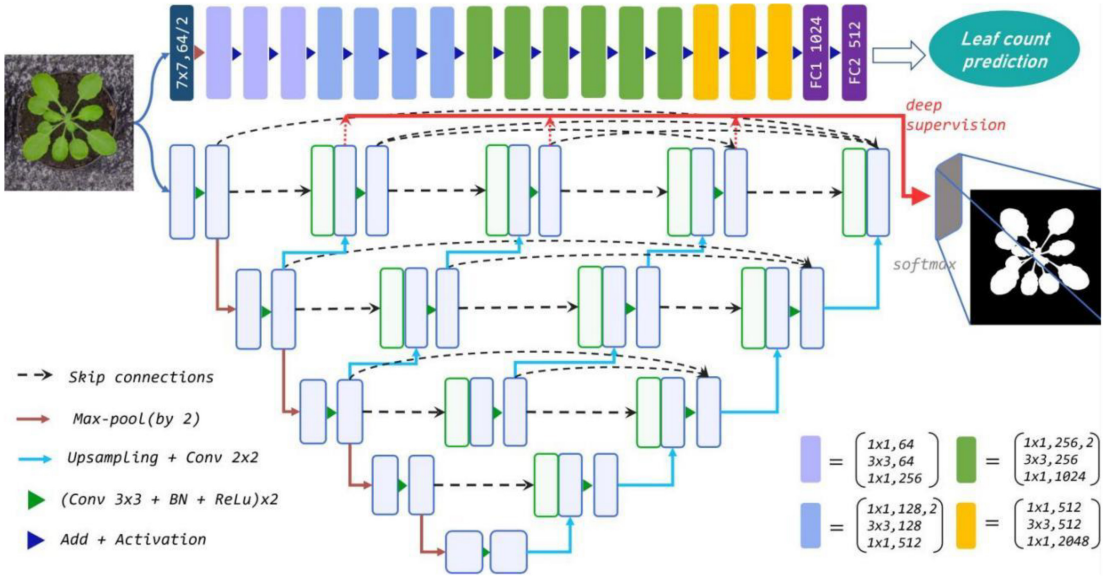


FIGURE 1 | The proposed parallel two-stream network combines leaf counting and segmentation tasks. Top row: the modified Resnet50 regression model for leaf counting with 16 residual blocks. Remaining rows: U-Net++ for segmentation *via* multi-use of the features from different semantic levels (layers). Each blue box corresponds to a multi-channel feature map, and the green boxes represent copied feature maps. The arrows denote various operations.

The binary cross-entropy and dice coefficient are combined to form the final loss function, which is defined as

$$\mathcal{L}(Y_{gt}, Y_{pred}) = -\frac{1}{N}\sum_{b=1}^{N}\left(\frac{1}{2}\cdot Y_{gt}^{b}\cdot \log Y_{pred}^{b} + \frac{2\cdot Y_{gt}^{b}\cdot Y_{pred}^{b}}{Y_{gt}^{b} + Y_{pred}^{b}}\right) \quad (3)$$

where Y_{gt}^{b} and Y_{pred}^{b} denote the predict map and ground truth map of b-th image, respectively, and N denotes the batch size.

The objective function takes the form of a logarithmic logic function as a replacement for the complex softmax multi-class prediction function. Forward propagation infers the prediction results and compares them with the true value annotations to generate cross-entropy loss. Backward propagation updates the model weight parameters. In this way, the task of plant segmentation is transformed into a binary classification problem that is suitable for plant segmentation. The re-designed skip pathways take effect on the output of the fused features and simplify the optimisation on the shallow, middle, and profound output results for varying degrees, *via* tuning the overall parameter of the network.

Learning Count Model With Segmentation

During leaf counting, the estimated number of leaves tends to exceed its ground truth. This is because the lower part of a leaf might be occluded by other leaves, or the leaves are incomplete and fragmentary on their own, which would be ignored by the counting model. To address this problem, we introduced the auxiliary cue, i.e., the segmented mask to guide the learning of the counting model. Also, it is widely acknowledged the counting model could fail due to the lacking of available samples belonging to certain class in the training dataset. The labelling for leaf counting is also time-consuming. Such data scarcity is often met in the data-driven methods such as deep learning. Thus, we augmented the samples by combining the segmented mask and the original images, which enhance the model to effectively capture the occluded leaves and the hardly detected leaves in plant image under the assistance of segmented binary mask.

Inspired by the work of He et al. (2016), we employed Resnet50 network as our backbone architecture due to its superb performance in image recognition. For our regression task, we modified the Resnet50 network by replacing the last layer with a fully connected layer with one-dimension output, which acts as a regression model for leaf counting. The modified network uses the combined samples from the segmentation mask and the original images as input, and applies convolution with a 7×7 filter followed by a series of convolutions, ending with fully connected layers to determine the number of plant predictions. Residual learning is also used to overcome the inefficient learning and the possibility of over-fitting due to deep network, where the skip connections resolve the degradation problem by taking the output of the previous layers as the input of the latter. For instance, when an input is x and the learned features are denoted as $H(x)$, then the residual learning features is $F(x) = H(x) - x$. The stacked-layer learns new features on top of the input features, and a residual unit is given by

$$y_l = h(x_l) + F(x_l, W_l), x_{l+1} = f(y_l) \quad (4)$$

where x_l and x_{l+1}, respectively, represent the input and output of the l th residual unit, and each residual unit contains multiple

layers of structure. F represents the learned residual block, $h(x_l) = x_l$ is the constant mapping, f is the ReLU activation function. Thus, the learned features from shallow l to deep L are

$$x_L = x_l + \sum_{i=l}^{L-1} F(x_i, W_i) \tag{5}$$

A chain rule is used to aid the reverse process of gradients, i.e.,

$$\frac{\partial \text{ loss}}{\partial x_l} = \frac{\partial \text{loss}}{\partial x_L} \cdot \frac{\partial x_L}{\partial x_l} = \frac{\partial \text{loss}}{\partial x_L} \cdot \left(1 + \frac{\partial}{x_L} \sum_{i=l}^{L-1} F(x_i, W_i)\right) \tag{6}$$

where $\frac{\partial \text{loss}}{\partial x_L}$ denotes the gradient of the loss function reaching L, the value 1 in the parentheses indicates that the shortcut connection mechanism propagates the gradient without loss, while other residual gradient passes through a layer with weights indirectly. In this context, 1 is selected to make the residual gradient easier to learn and thus avoid the gradient vanishing.

To better train the regression model, we employed mean squared error (MSE) as the loss function. Given an image i and the ground truth leaf count $y^i_{gt,c}$, the loss function L_c is determined by

$$L_c = \frac{1}{m} \sum_{i=1}^{m} \left(y^i_{pred,c} - y^i_{gt,c}\right)^2 \tag{7}$$

where m is the image number and $y^i_{pred,c}$ denotes the predicted leaf count.

With respect to our regression task, the last fully-connected layer with 1,000 neurons initially used for classification is replaced by a layer with a single neuron, which allows for the output estimation of leaf number. The neuron is to regress the correct leaf numbers given the input images. To obtain the rich prior knowledge, the regression network is pre-trained on ImageNet for parameter initialization, and then fine-tuned on the used datasets. Our regression model is shown in the top row of **Figure 1**. Note that the combination of segmentation and RGB images extends the input channel from 3 to 4. By extending the channel, an additional binary channel is added to the leaf count regression model to convey pure semantic information of leaf and suppress bias from features in the background of the training images, e.g., the soil, moss, pot, etc., that differ between datasets. At the same time, the RGB channels enable the network to retain the rich local texture and context information that the binary mask fails to capture, thus enhancing the robustness of our model. In addition, our regression model does not require any bounding box or centre point annotation, which can be efficiently applied to deal with more complex scenes.

U-Net remains the preferred choice for the maintenance of fine edge binary segmentation. The design of skip connections greatly enriches the information received by the decoder, and *via* specially trained end-to-end, U-Net performs high-precision segmentation for small training samples. When applied in leaf segmentation, the architecture extracts the edge details, size, and shape diversity in the low-level information and uncovers the discriminative high-level information of the target leaf. This advantage reduces the overall size of the dataset required for training. Furthermore, due to the effective reuse of extracted features and an ability to capture the targets, the architecture achieves an implicit data argumentation and speeds up the convergence for the binary tasks during training.

However, since the leaf dataset (with sub-datasets A1–A4) varies in the degree of occlusion, leaf numbers and leaf size, we only combined the same-scale information not previously countered. Designing U-net with different depth for each layer may be an idea but such an approach has not been widely applied. To address this, we adopt U-Net++ (remaining rows of **Figure 1**) as the feature extractor for segmentation, which extends U-Net with denser cross-layer concatenation and shortens the semantic gap between the encoder and decoder by fusing spatial information from shallow to deep cross layers. The architecture makes full use of contextual features and semantic information from the same dimension, and it captures the detailed features of the target. Moreover, using the pruning scheme basing on the module which receives the best estimation during training, the network is adjustable and customisable. For instance, it is customised to the most suitable size and saves unnecessary storage space. This is equivalent to the maintenance of any useful feature we acquired and the distinctive design for each dataset in one end-to-end network.

EXPERIMENTS

We thoroughly assess the effectiveness of our proposed framework on the widely used plant phenotyping dataset including its four sub-datasets (see Section "Dataset and Data Pre-processing"). We conducted extensive experiments on both plant segmentation and leaf counting, and compared the performance of our method with the state-of-the-art methods for validation. We explored three segmentation architectures using three different backbone networks, i.e., MobileNet, ResNet, and VGGNet on the four sub-datasets, and compared our method with the state-of-the-art leaf segmentation methods. We also performed the experiments to demonstrate the effectiveness of the proposed leaf counting method, comparing it with the state-of-the-art leaf counting methods.

Dataset and Data Pre-processing

The dataset used in our experiments belongs to the Leaf Segmentation and Counting Challenge (LCC and LSC) held as part of the Computer Vision Problems in Plant Phenotyping (CVPPP 2017) workshop (Giuffrida et al., 2015). The dataset is divided into training set and testing set, which consists of 810 and 275 top-down view RGB images of either Tobacco or Arabidopsis plants, respectively. Both training and testing images are grouped into four folders, i.e., four sub-datasets which vary from the species and means of collection such as imaging setups and labs. The training sets include 128, 31, 27, 624 images and the testing sets contain 33, 9, 65, 168 images for A1, A2, A3, and A4 respectively. The sub-datasets A1 and A2 include Arabidopsis images collected from growth

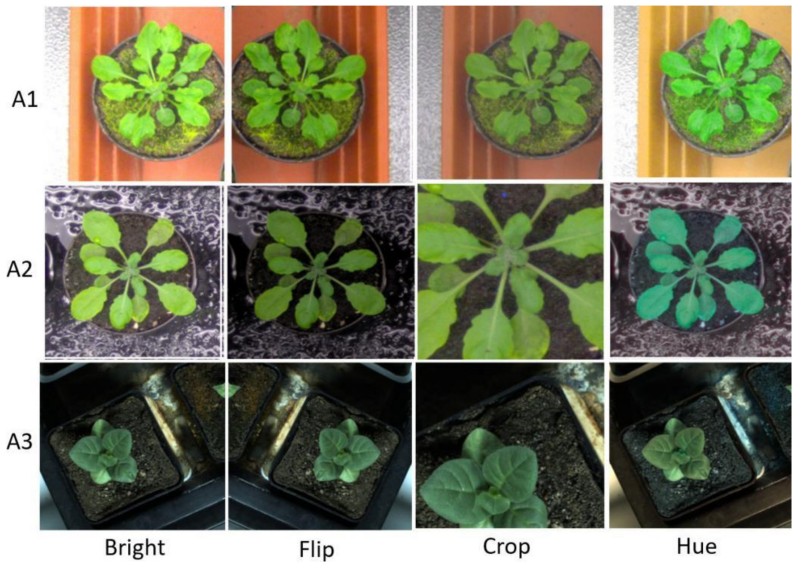


FIGURE 2 | Augmentation samples for training the segmentation network to avoid the risk of over-fitting.

chamber experiments with different field of views covering many plants and then cropped to a single plant image with the size of approximately 500 × 500 pixels. Sub-dataset A3 contains tobacco images at 2,000 × 2,500 pixels with the field of view chosen to encompass a single plant. Sub-dataset A4 is a subset of another public Arabidopsis dataset. The dataset provides the corresponding annotations in binary segmentation with 1 and 0, respectively, denoting plant and background pixels. All the folders contain the ground truth binary mask used for whole plant segmentation (i.e., semantic segmentation). For the experiment of plant segmentation, we follow the training strategy from Aich and Stavness (2017), and also use the combination of all sub-datasets (referred as to *All*) for training to achieve more robust model.

In our work, we addressed two problems caused by a dataset as follows: (1) Deep learning based methods require a huge amount of training samples while the availability of the dataset of plant leaf with annotations is limited, causing data scarcity; and (2) Small and overlapping leaf instances brought a challenge for plant segmentation and leaf counting. Data augmentation is a widely used technique in deep learning to increase the number of samples and provide more diversity to the deep neural networks. In this context, we also employed data augmentation to address the above two problems.

Moreover, we first reshaped the size of training images to 480 × 480 pixels and normalized. Following the resize operation, we conducted the following scheme for data augmentation: (1) Random-Rotate with an interval of 90 to increase the network invariance to slight angular changes; (2) Flip: horizontal, vertical, and horizontal+ vertical; (3) Resize the images to increase the network invariance to different image resolutions; (4) Gamma transform to extend the data by changing the image greyscale; (5) Random-Brightness: the clarity of object depends on scene lighting and camera sensitivity, thus random changing the image brightness improves the illumination invariance of the network; (6) Random change in the contrast range to increase the network invariance to shadows and improve the network performance in low light conditions; (7) Hue Saturation Brightness (HSV): changes in colour channels, degree of lightness or darkness of a colour; and (8) Normalise a characteristic linear transformation which scales a specific range of data values retaining the original data distribution. Selected augmentation processes are shown in **Figure 2**.

Implementation Details and Evaluation Protocol

All images from the training set are randomly split into two sets for training and validation with the split ratio of 0.8 and 0.2, respectively. Images from the testing set are used for evaluating the segmentation performance. We used the validation set to verify the hyper-parameters (see **Table 1**) during the training of the initial experiments.

Network Parameter Setting

All our experiments are performed on the PyTorch platform with NVIDIA 2080Ti GPU. We used the data augmentation to increase the number of samples as in Section "Dataset and Data Pre-processing." This module contributes to preventing over-fitting for the relatively small plant datasets and ensure the

TABLE 1 | Hyper-parameters used for training.

Epochs	100
Batch-size	4
Optimizer	Adam
Learning-rate	1e-3
Weight-decay	1e-4
Factor	0.1

model produces promising results when segmenting on new data *via* learning multiple variations (Holmberg, 2020). The binary mask is transformed the same way, to maintain the consistency between images and annotations (except for the transformation regarding colours).

We randomly sampled four samples to form a mini-batch with batch size of four to guarantee the convergence of training. Adam is adopted as the optimizer for its fast convergence rate to train the model for a total of 100 epochs, where the results remain stable with no further improvement. The weight decay factor is set to 0.0001 and the learning rate is constantly set as 0.001.

Metrics for Segmentation

We employed the intersection of union (IoU) as the evaluation metric, which is widely used in segmentation. IoU is used to determine the spatial overlap between the segmented leaf region and its ground truth, i.e.,

$$\text{IoU}(\%) = \frac{|P_{\text{gt}} \cap P_{\text{pred}}|}{|P_{\text{gt}}| + |P_{\text{pred}}|} \tag{8}$$

where P_{gt} and P_{pred}, respectively, denote the ground truth mask and the prediction mask. Due to the problem of class imbalance between positive and negative samples, it is insufficient to use accuracy as evaluation metric. For better evaluation, we introduced two more metrics: Precision and Recall. Precision is used to determine the portion of segmented leaf region pixels that matches with the ground truth, i.e.,

$$\text{Precision}(\%) = \frac{TP}{TP + FP} \times 100 \tag{9}$$

Recall is used to determine the portion of ground-truth pixels present in the segmented leaf region, i.e.,

$$\text{Recall}(\%) = \frac{TP}{TP + FN} \times 100 \tag{10}$$

where True Positive (TP), False Negative (FN), and False Positive (FP) respectively denote the number of leaf region pixels correctly identified, the number of leaf region pixels unidentified, and the number of leaf region pixels falsely identified.

Metrics for Count

To evaluate how good a leaf count method is in estimating the correct number of leaves, we employed the regression metrics: Difference in Count (DiC), Absolute Difference in Count (ADiC), and mean squared error (MSE) calculated as follows:

$$\text{DiC} = \frac{1}{m}\sum_{i=1}^{m}\left(y_{\text{gt,c}}^{(i)} - y_{\text{pred,c}}^{(i)}\right) \tag{11}$$

$$\text{ADiC} = \frac{1}{m}\sum_{i=1}^{m}\left|\left(y_{\text{gt,c}}^{(i)} - y_{\text{pred,c}}^{(i)}\right)\right| \tag{12}$$

$$\text{MSE} = \frac{1}{m}\sum_{i=1}^{m}\left(y_{\text{gt,c}}^{(i)} - y_{\text{pred,c}}^{(i)}\right)^2 \tag{13}$$

Experimental Analysis

Segmentation Analysis

In the first experiment, we evaluated the effectiveness of our segmentation model on plant images by using different segmentation architectures and backbones for comparison. FCN8, PSPNet, U-Net are selected as the basic encoder and decoder architectures, where ResNet and VGG are used as backbones due to its good ability of depicting 2D images. The

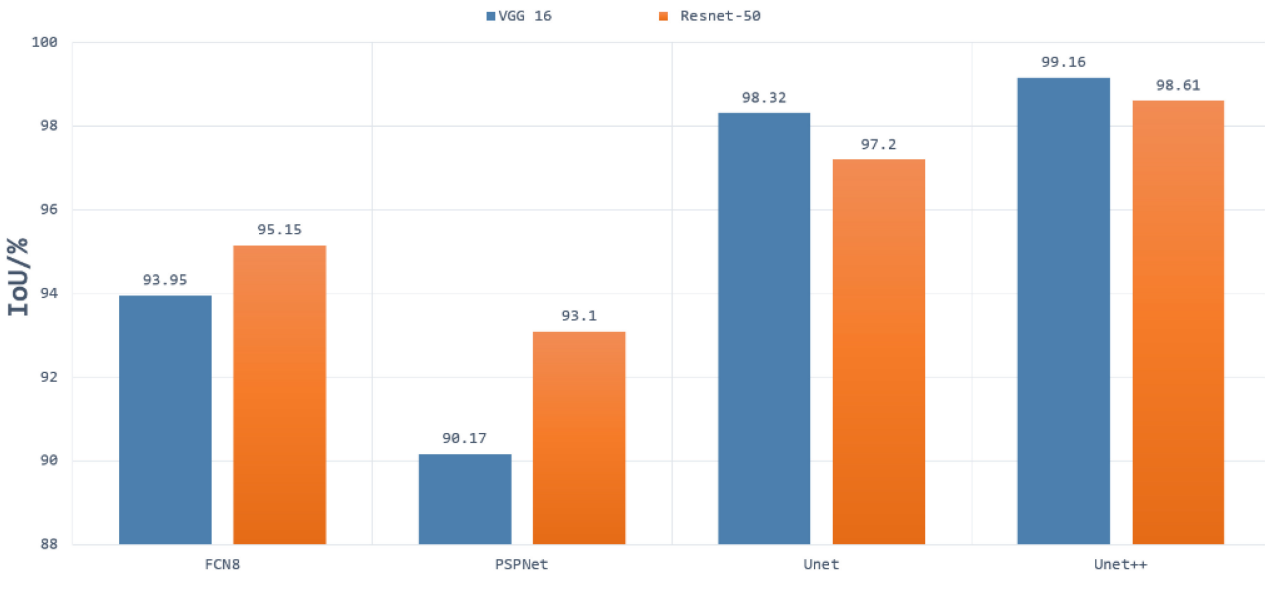


FIGURE 3 | Results of segmentation using Resnet50 and VGG16 as backbone in FCN, PSPnet, U-Net, and U-Net++ architectures.

TABLE 2 | Segmentation results on each sub-dataset and their com- bination using different basic architectures.

IoU (%)	All	A1	A2	A3	A4
FCN	93.95	93.45	89.17	88.51	92.23
PSPNet	90.17	94.34	90.55	91.19	93.83
U-Net	98.32	98.51	97.76	94.72	97.17
U-Net++	99.11	98.29	97.98	95.90	97.23

comparative segmentation performance in terms of IoU on the combination of all sub-datasets are provided in **Figure 3**. It is evident from **Figure 3** that the segmentation results generated by our segmentation model outperforms the other architectures. Among different models, using VGG as backbone performs constantly better than using ResNet as backbone. To evaluate the performance of dealing with a variety of scenes, we evaluated our model on the four individual sub-datasets and the results are shown in **Table 2**. The U-Net++ performs significantly better than the state-of-the-art segmentation methods. For better illustration, the segmentation results for images in sub- dataset A1 using different models together with ground truth are shown in **Figure 4**. Although all the three semantic segmentation methods can obtain clear segmentation results on A1, the U-Net++ retains the boundary and detail information. For the relative scarce sub-dataset A3 which only contains 27 tobacco images, the proposed method still shows a stable IoU. For each sub-dataset, the network generates segmentation results that are almost consistent with the corresponding binary template, from both quantitative and qualitative standpoints.

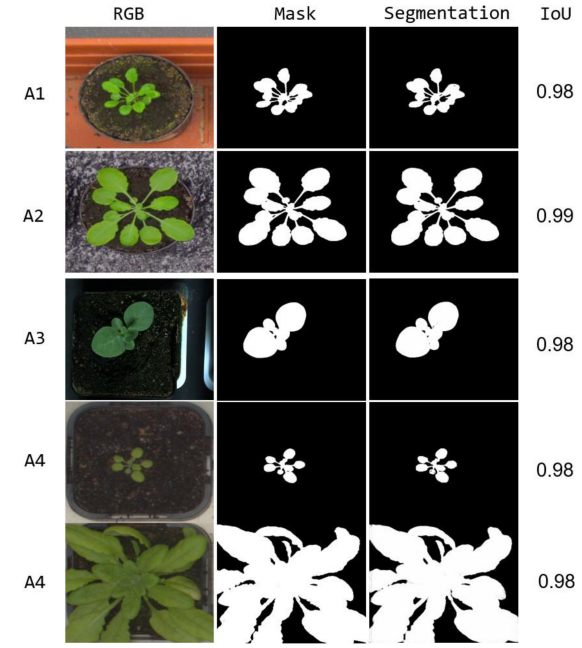


FIGURE 5 | Segmentation result for each sub-dataset, with the corresponding IoU provided at the right.

During the training for segmentation, the sigmoid function produces outputs in the range [0 ... 1]. While calculating the loss, greater weight is assigned for the boundary pixels. The weight

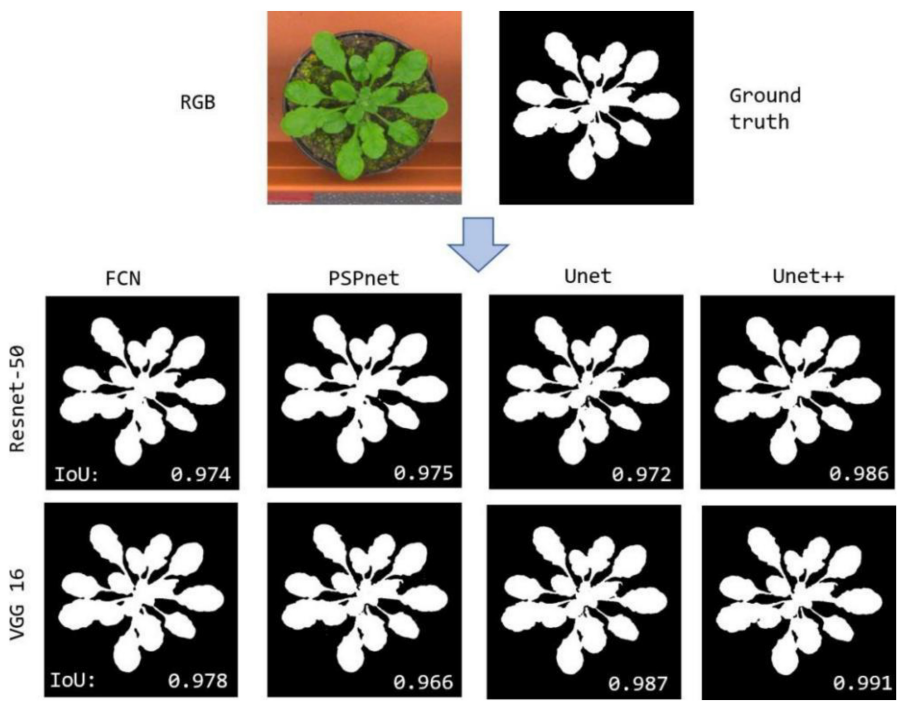


FIGURE 4 | Comparing segmentation results on the same RGB image.

map is then calculated using

$$w(\mathbf{x}) = w_c(\mathbf{x}) + w_0 \cdot \exp\left(-\frac{\left(d_1(\mathbf{x}) + \right)d_2(\mathbf{x})\right)^2}{2\sigma^2}\right) \quad (14)$$

where $w_c(x)$ is the category weight based on the frequency of occurrence of each category in the training dataset; $d_1(x)$ represents the distance between the object pixel and the nearest boundary. $d_2(x)$ represents the same distance for the second nearest boundary. In our work, we set the threshold σ to 0.5 to obtain the segmentation weight map. The segmentation results using our method on different sub-datasets are shown in **Figure 5**. Our model generates the segmentation results that are almost coincident visually with the ground-truth mask for each sub-dataset. For A3 sub-dataset which only contains 27 tobacco images with small leaf area, our method still shows a stable segmentation result. The results show our method effectively addresses segmentation under various scenes, i.e., with occlusions, small leaf area, and large leaf area, demonstrating good robustness.

We also compared the convergence rate of different segmentation models. The curves of the precision, recall, training cross entropy (CE) loss, and IoU are shown in **Figure 6**. The figure shows that all four networks selecting VGG16 as the encoder for feature extraction achieve good IoU scores consistently. In addition, **Figure 7** visualises the feature extraction process of our method using UNet++ with VGG from the early to late epochs. The process of feature extraction is smoother and faster to reach the convergence, which shows VGG can capture the meaningful representations for leaf images.

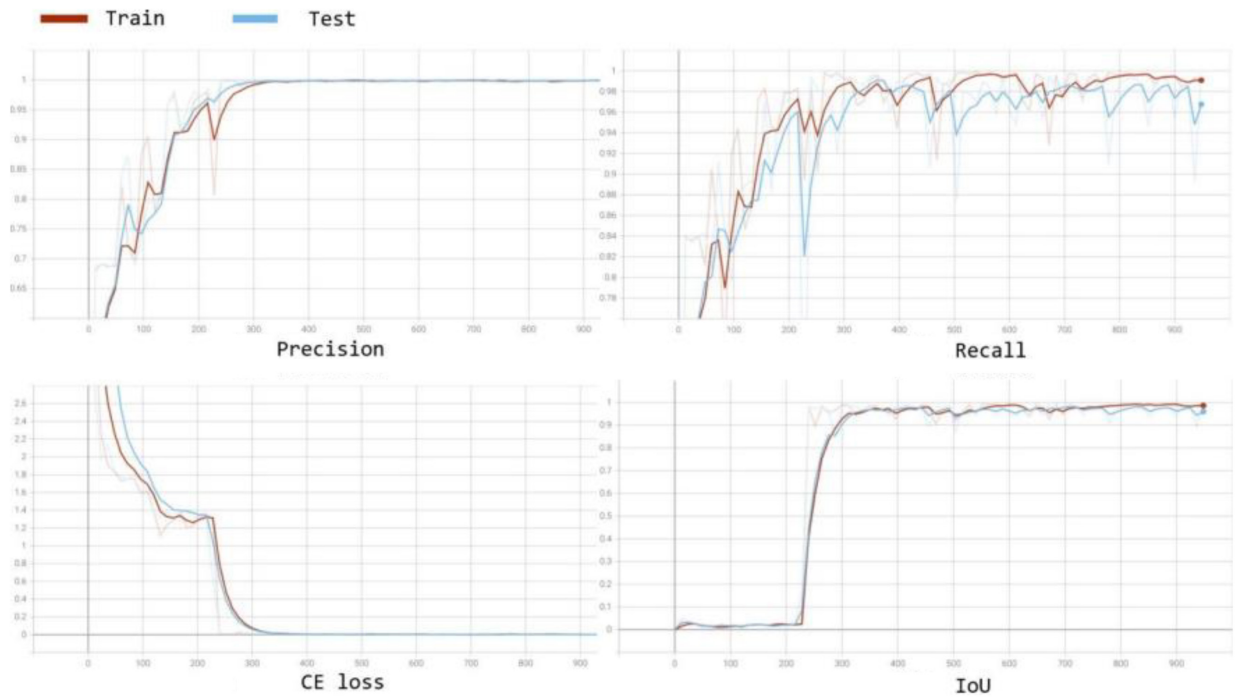


FIGURE 6 | Convergence curves for accuracy, loss, and IoU score on the validation set during the training process for comparison in terms of accuracy and convergence rate.

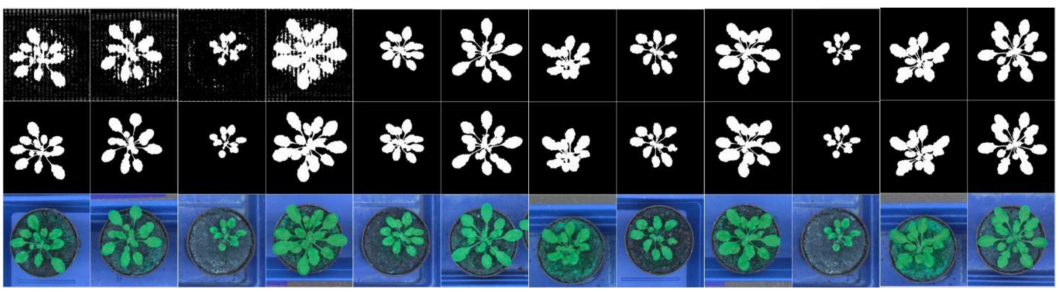


FIGURE 7 | Visualization for the feature extraction process of our method, arranged by time series from the early to late epochs. The first to third line images respectively show the predicted images, ground truth images and transformed RGB images.

TABLE 3 | Segmentation results on each sub-dataset and their combination using different basic architectures.

	SRGB					Ours				
	All	A1	A2	A3	A4	All	A1	A2	A3	A4
Precision	0.92	0.98	0.94	0.80	0.96	0.99	0.99	0.99	0.99	0.99
Recall	0.97	0.99	0.99	0.94	0.98	0.99	0.98	0.99	0.99	0.99
IoU	–	–	–	–	–	0.98	0.98	0.99	0.98	0.98

TABLE 4 | Counting results using different backbones with or without the auxiliary binary mask on CVPPP 2017 dataset (Bold values denote the best performance).

Metric	DiC	ADiC	MSE
Mobilenet			
RGB	–0.30	0.66	0.98
RGB+SBM	0.13	0.46	0.64
InceptionNet			
Rgb	0.29	0.61	1.20
RGB+SBM	0.20	0.43	0.54
VGGNet			
RGB	0.20	0.79	1.44
RGB+SBM	–0.12	0.37	0.44
Resnet50			
RGB	–0.12	0.60	0.89
RGB+SBM	**0.11**	**0.36**	0.42

For DiC, ADiC, and MSE, a lower value is better.

We compared the proposed segmentation model with the other state-of-the-art method that performed the experiment on plant (foreground) segmentation, i.e., SRGB (Aich and Stavness, 2017) using three metrics, i.e., Precision, Recall, and IoU and the results are shown in **Table 3**. Our method outperforms the SRGB method on two metrics, achieving the high performance on IoU. The results suggest that our approach is very effective for plant segmentation task in plant phenotyping.

Leaf Count Evaluations

In the second experiment, we evaluated the effectiveness of the proposed leaf counting method using segmented binary mask

TABLE 5 | Comparative evaluation of the proposed counting model with state-of-the-art methods.

	DiC	ADiC	MSE
IPK	–1.9 (2.7)	2.4 (2.1)	–
GLC	–0.51 (2.02)	1.43 (1.51)	4.31
Nottingham	–2.4 (2.8)	2.9 (2.3)	–
MSU	–2.3(1.8)	2.4 (1.7)	–
Wageningen	1.5 (4.4)	2.5 (3.9)	–
Proposed RGB+SBM	0.11 (0.98)–	0.36 (0.93)	0.42

(referred as RGB+SBM). During the experiment, the number of input channels must be consistent with the input size of the backbone models, i.e., 3 channels. In this way, when a binary image with single channel is fed into the model, the values of the single channel are extended to three channels by duplication, forming an image with three channels. The resulting three-channel images are mixed with the RGB image samples to increase the number of training samples, facilitating the stability of leaf counting. To validate the effectiveness of our counting model for leaf counting, we adopted different backbones for our leaf counting task, e.g., MobileNet, VGGNet, InceptionNet, and ResNet, and report the results in **Table 4**. Moreover, to further explore the potential benefit of the auxiliary binary mask, we conducted an ablation experiment on with/without using the binary channel, and the result is also shown in **Table 4**. In **Table 4**, RGB denotes the method without using the binary mask, while RGB+SBM denotes that our method using the auxiliary binary mask. It is observed from the table that the count model using the ResNet50 backbone performs the best among the backbones. The binary mask increases the count performance in all metrics, where the MSE drops from 0.89 to 0.04, DiC from 0.02 to 0.01, and ADiC from 0.60 to 0.36. These results validate our assumption that binary mask improves the accuracy and robustness for the leaf count model, due to its capability to deal with background interferences.

We used the scatter diagram to visually illustrate the correlation between the estimated leaf numbers and their ground truth, and the results are shown in **Figure 8**, which is also

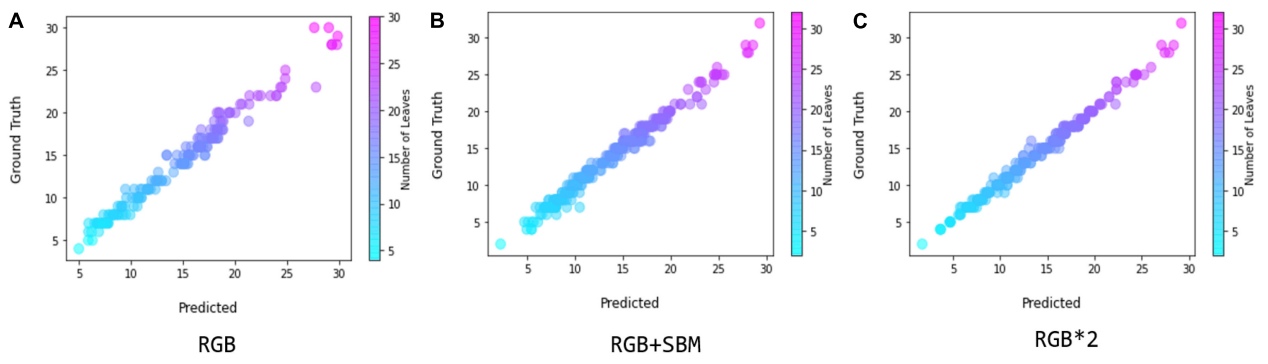


FIGURE 8 | Comparison between the coefficient of determination in the implementation of scatter graphics, where **(A)** denotes using only RGB image, **(B)** denotes using the mixture of RGB and segmented binary mask, and **(C)** denotes using the double RGB images by simple copy.

for the evaluation of our regression model. The higher overlap between the scatter plots of estimation and the ground truth indicates a better agreement. **Figure 8** shows that the binary mask significantly enhances the agreement between the ground truth and the estimation, as the error distribution in leaf count is constantly confined within smaller region. If directly doubling the number of the input samples by simple copy, referred as RGB *2, we find that the performance is almost the same as with the mixture of RGB and binary mask images. In the experiments, the time cost using double RGB images is higher than using the combination of RGB and binary mask images. Thus, we conclude that using the auxiliary binary mask to guide the leaf counting is a simple but effective way for improving the performance of counting.

In addition, we reported the quantitative comparison of our leaf counting method with state-of-the-art methods i.e., GLC (Giuffrida et al., 2015), IPK (Pape and Klukas, 2015), Nottingham (Scharr et al., 2016), MSU (Scharr et al., 2016), and Wageningen (Scharr et al., 2016), as shown in **Table 5**. For fair comparison, we used A1, A2, A3 from testing set for testing the counting performance. Overall, the proposed leaf counting model using segmented binary mask achieves the best performance with lower values in the metrics of DiC, ADiC, and MSE. This shows the proposed counting model estimates the number of leaves with adequate accuracy and stability.

CONCLUSION

In this study, we focus on dealing with two fundamental tasks in plant phenotyping, i.e., plant segmentation and leaf counting, and propose a two-stream deep learning framework for automatic segmenting and counting leaves with various size and shape from two-dimensional plant images. In the first stream, a multi-scale segmentation model using spatial pyramid is developed to extract the whole plant in different size and shape, where the fine-grained details of leaves are captured using deep feature extractor. In the second stream, a regression counting model is proposed to estimate the number of leaves without any pre-detection, where the auxiliary binary mask is introduced to enhance the counting performance by effectively alleviating the influence of complex background. Extensive experiments on a publicly available plant phenotyping dataset show that the proposed framework achieves a promising performance both in the task of plant segmentation and leaf counting, providing a reference for the automatic analysis of plant. Future work will focus in increasing the robustness of the tasks of segmentation and the counting to deal with varying environments.

DATA AVAILABILITY STATEMENT

Publicly available datasets were analyzed in this study. This data can be found here: https://www.plant-phenotyping.org/CVPPP2017.

AUTHOR CONTRIBUTIONS

XF contributed to writing the draft and designing the ideas. RZ contributed to conducting experiments. TT contributed to editing the draft. SD and QY contributed to algorithm supervision. All authors contributed to the article and approved the submitted version.

REFERENCES

Aich, S., and Stavness, I. (2017). "Leaf counting with deep convolutional and deconvolutional networks," in *Proceedings of the IEEE International Conference on Computer Vision Workshops*, Washington, DC, 2080–2089.

Bell, J., and Dee, H. M. (2019). Leaf segmentation through the classification of edges. *arXiv [Preprint]* doi: 10.48550/arXiv.1904.03124

Chen, L.-C., Papandreou, G., Kokkinos, I., Murphy, K., and Yuille, A. L. (2017). Deeplab: semantic image segmentation with deep convolutional nets, atrous convolution, and fully connected crfs. *IEEE Trans. Pattern Anal. Mach. Intell.* 40, 834–848. doi: 10.1109/TPAMI.2017.2699184

Choudhury, S. D., Samal, A., and Awada, T. (2019). Leveraging image analysis for high-throughput plant phenotyping. *Front. Plant Sci.* 10:508. doi: 10.3389/fpls.2019.00508

da Silva, N. B., and Goncalves, W. N. (2019). "Regression in convolutional neural networks applied to plant leaf counting," in *Proceedings of the Anais do XV Workshop de Visão Computacional, SBC*, São Bernardo do Campo, 49–54.

Dhaka, V. S., Meena, S. V., Rani, G., Sinwar, D., Ijaz, K. M. F., Woźniak, M., et al. (2021). A survey of deep convolutional neural networks applied for prediction of plant leaf diseases. *Sensors* 21:4749. doi: 10.3390/s21144749

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

Dobrescu, A., Valerio Giuffrida, M., and Tsaftaris, S. A. (2017). "Leveraging multiple datasets for deep leaf counting," in *Proceedings of the IEEE International Conference on Computer Vision Workshops*, Venice, 2072–2079.

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

Giuffrida, M. V., Doerner, P., and Tsaftaris, S. A. (2018). Pheno-deep counter: a unified and versatile deep learning architecture for leaf counting. *Plant J.* 96, 880–890. doi: 10.1111/tpj.14064

Giuffrida, M. V., Minervini, M., and Tsaftaris, S. (2015). "Learning to count leaves in rosette plants," in *Proceedings of the Computer Vision Problems in Plant Phenotyping (CVPPP)*, eds H. S. S. A. Tsaftaris and T. Pridmore (Swansea: BMVA Press), 1.1–1.13. doi: 10.5244/C.29.CVPPP.1

Gomes, D. P. S., and Zheng, L. (2020). Leaf segmentation and counting with deep learning: on model certainty, test-time augmentation, trade-offs. *arXiv [Preprint]* doi: 10.48550/arXiv.2012.11486

He, K., Gkioxari, G., Dollár, P., and Girshick, R. (2017). "Mask r-cnn," in *Proceedings of the IEEE International Conference on Computer Vision*, Venice, 2961–2969.

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*, Las Vegas, NV, 770–778.

Holmberg, J. (2020). *Targeting the Zebrafish Eye using Deep Learning-Based Image Segmentation*. Uppsala: Uppsala University Publications.

Itzhaky, Y., Farjon, G., Khoroshevsky, F., Shpigler, A., and Bar-Hillel, A. (2018). *Leaf Counting: Multiple Scale Regression and Detection using Deep CNNS*. Beer Sheva: Ben Gurion University of the Negev, 328.

Kong, Y., Li, H., Ren, Y., Genchev, G. Z., Wang, X., Zhao, H., et al. (2020). Automated yeast cells segmentation and counting using a parallel u-net based two-stage framework. *OSA Continuum* 3, 982–992. doi: 10.1364/osac.388082

Kumar, J. P., and Domnic, S. (2019). Image based leaf segmentation and counting in rosette plants. *Inform. Process. Agric.* 6, 233–246. doi: 10.1016/j.inpa.2018.09.005

Kumar, J. P., and Domnic, S. (2020). Rosette plant segmentation with leaf count using orthogonal transform and deep convolutional neural network. *Mach. Vis. Appl.* 31, 1–14.

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

Kuznichov, D., Zvirin, A., Honen, Y., and Kimmel, R. (2019). "Data augmentation for leaf segmentation and counting tasks in rosette plants," in *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops*, Long Beach, CA.

Lecun, Y., Bengio, Y., and Hinton, G. (2015). Deep learning. *Nature* 521, 436–444.

Lin, Z., and Guo, W. (2020). Sorghum panicle detection and counting using unmanned aerial system images and deep learning. *Front. Plant Sci.* 11:534853. doi: 10.3389/fpls.2020.534853

Long, J., Shelhamer, E., and Darrell, T. (2015). "Fully convolutional networks for semantic segmentation," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, Boston, MA, 3431–3440.

Lu, H., and Cao, Z. (2020). Tasselnetv2+: a fast implementation for high-throughput plant counting from high-resolution rgb imagery. *Front. Plant Sci.* 11:541960. doi: 10.3389/fpls.2020.541960

Minervini, M., Scharr, H., and Tsaftaris, S. A. (2015). Image analysis: the new bottleneck in plant phenotyping [applications corner]. *IEEE Signal Process. Mag.* 32, 126–131. doi: 10.1109/msp.2015.2405111

Mishra, P., Sadeh, R., Bino, E., Polder, G., Boer, M. P., Rutledge, D. N., et al. (2021). Complementary chemometrics and deep learning for semantic segmentation of tall and wide visible and near-infrared spectral images of plants. *Comput. Electron. Agric.* 186:106226. doi: 10.1016/j.compag.2021.106226

Montero, F., De Juan, J., Cuesta, A., and Brasa, A. (2000). Nondestructive methods to estimate leaf area in vitis vinifera l. *Hort Sci.* 35, 696–698. doi: 10.1093/aob/mcf059

Rahnemoonfar, M., and Sheppard, C. (2017). Deep count: fruit counting based on deep simulated learning. *Sensors* 17:905. doi: 10.3390/s17040905

Ren, M., and Zemel, R. S. (2017). "End-to-end instance segmentation with recurrent attention," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, Honolulu, HI, 6656–6664.

Romera-Paredes, B., and Torr, P. H. S. (2016). "Recurrent instance segmentation," in *Proceedings of the European Conference on Computer Vision*, (Cham: Springer), 312–329. doi: 10.1007/978-3-319-46466-4_19

Ronneberger, O., Fischer, P., and Brox, T. (2015). "U-net: convolutional networks for biomedical image segmentation," in *International Conference on Medical Image Computing and Computer-Assisted Intervention* (Berlin: Springer), 234–241. doi: 10.1007/978-3-319-24574-4_28

Scharr, H., Minervini, M., French, A., Klukas, P. C., Kramer, D. M., Liu, X., et al. (2016). Leaf segmentation in plant phenotyping: a collation study. *Mach. Vis. Appl.* 27, 585–606. doi: 10.1007/s00138-015-0737-3

Siebner, H., Callicott, J., Sommer, T., and Mattay, V. (2009). From the genome to the phenome and back: linking genes with human brain function and structure using genetically informed neuroimaging. *Neuroscience* 164, 1–6. doi: 10.1016/j.neuroscience.2009.09.009

Tassis, L. M., de Souza, J. E. T., and Krohling, R. A. (2021). A deep learning approach combining instance and semantic segmentation to identify diseases and pests of coffee leaves from in-field images. *Comput. Electron. Agric.* 186:106191. doi: 10.1016/j.compag.2021.106191

Tsaftaris, S. A., Minervini, M., and Scharr, H. (2016). Machine learning for plant phenotyping needs image processing. *Trends Plant Sci.* 21, 989–991. doi: 10.1016/j.tplants.2016.10.002

Ubbens, J., Cieslak, M., Prusinkiewicz, P., and Stavness, I. (2018). The use of plant models in deep learning: an application to leaf counting in rosette plants. *Plant Methods* 14, 1–10. doi: 10.1186/s13007-018-0273-z

Ward, D., Moghadam, P., and Hudson, N. (2018). Deep leaf segmentation using synthetic data. *arXiv [Preprint]* doi: 10.48550/arXiv.1807.10931

Yang, W., Feng, H., Zhang, X., Zhang, J., Doonan, J. H., Batchelor, W. D., et al. (2020). Crop phenomics and high-throughput phenotyping: past decades, current challenges, and future perspectives. *Mol. Plant* 13, 187–214. doi: 10.1016/j.molp.2020.01.008

Zhao, H., Shi, J., Qi, X., Wang, X., and Jia, J. (2017). "Pyramid scene parsing network," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, San Juan, PR, 2881–2890.

Zhou, Z., Siddiquee, M. M. R., Tajbakhsh, N., and Liang, J. (2018). "Unet++: a nested u-net architecture for medical image segmentation," in *Deep Learning in Medical Image Analysis and Multimodal Learning for Clinical Decision Support*, eds L. Maier-Hein, T. Syeda-Mahmood, Z. Taylor, Z. Lu, D. Stoyanov, A. Madabhushi, et al. (Cham: Springer), 3–11. doi: 10.1007/978-3-030-00889-5_1

Zhu, Y., Aoun, M., Krijn, M., Vanschoren, J., and Campus, H. T. (2018). "Data augmentation using conditional generative adversarial networks for leaf counting in Arabidopsis plants," in *Proceedings of the The British Machine Vision Conference (BMVC)*, Newcastle, 324.

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: 24 May 2022
doi: 10.3389/fpls.2022.862558

An Algorithm for Precipitation Correction in Flood Season Based on Dendritic Neural Network

*Tao Li[1], Chenwei Qiao[1], Lina Wang[1], Jie Chen[2] and Yongjun Ren[3]**

[1] School of Artificial Intelligence, Nanjing University of Information Science and Technology, Nanjing, China, [2] School of Electronics and Information Engineering, Nanjing University of Information Science and Technology, Nanjing, China, [3] School of Computer Software, Nanjing University of Information Science and Technology, Nanjing, China

In recent years, the National Climate Center has developed a dynamic downscaling prediction technology based on the Climate-Weather Research and Forecasting (CWRF) regional climate model and used it for summer precipitation prediction, but there are certain deviations, and it is difficult to predict more accurately. The CWRF model simulates the summer precipitation forecast data from 1996 to 2019 and uses a combination of dendrite net (DD) and artificial neural networks (ANNs) to conduct a comparative analysis of summer precipitation correction techniques. While summarizing the characteristics and current situation of summer precipitation in the whole country, the meteorological elements related to precipitation are analyzed. CWRF is used to simulate summer precipitation and actual observation precipitation data to establish a model to correct the precipitation. By comparing with the measured data of the ground station after quality control, the relevant evaluation index analysis is used to determine the best revised model. The results show that the correction effect based on the dendritic neural network algorithm is better than the CWRF historical return, in which, the anomaly correlation coefficient (ACC) and the temporal correlation coefficient (TCC) both increased by 0.1, the mean square error (MSE) dropped by about 26%, and the overall trend anomaly (Ps) test score was also improved, showing that the machine learning algorithms can correct the summer precipitation in the CWRF regional climate model to a certain extent and improve the accuracy of weather forecasts.

Keywords: machine learning, summer precipitation correction, dendrite net (DD) and artificial neural networks (ANN), mean square error (MSE), temporal correlation coefficient (TCC), anomaly correlation coefficient (ACC)

OPEN ACCESS

Edited by:
Lei Shu,
Nanjing Agricultural University, China

Reviewed by:
Margarita Gamarra,
Universidad de la Costa, Colombia
Xiaoqin Wei,
North Sichuan Medical College, China
Ying Zhao,
Henan Agricultural University, China
Chuanlei Zhang,
Tianjin University of Science and Technology, China

****Correspondence:***
Yongjun Ren
renyj100@126.com

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

Received: *26 January 2022*
Accepted: *06 April 2022*
Published: *24 May 2022*

Citation:
Li T, Qiao C, Wang L, Chen J and Ren Y (2022) An Algorithm for Precipitation Correction in Flood Season Based on Dendritic Neural Network. Front. Plant Sci. 13:862558. doi: 10.3389/fpls.2022.862558

INTRODUCTION

Climate prediction is the process of predicting the likely trend of climate development in the future based on the changing laws of the past climate. In recent years, the forward-looking role of climate prediction in disaster prevention and mitigation has been recognized more and more, and the demand for prediction in all walks of life is increasing. With the needs of social and economic development, research on climate prediction needs to be improved urgently (Yingdong et al., 2011; Bannister et al., 2019; Xinmin et al., 2019). Accurate precipitation data are essential for understanding climate change and associated hydrological responses from small

basins to large regions around the world (Pan et al., 2016). At present, global and regional climate models are the primary tools for climate change simulation and prediction research (Rummukainen, 2010), but they are restricted by the complexity of the climate system and the level of scientific development. Compared with actual observations, climate model simulations always show deviations in precipitation (Ren and Li, 2007; Kim et al., 2015, 2020).

Precipitation correction is an effective way to improve model forecasts. The concept of precipitation correction is proposed because there is some model data for flood season precipitation forecast in the current climate forecast business, but there are certain deviations. Therefore, it is hoped that the correction will reduce the error and improve the precipitation forecast accuracy and performance (Yao et al., 2017). With the development of weather forecasting technology, artificial intelligence, and data mining research, the use of intelligent computing and data mining technology to correct regional precipitation provides a new and effective method for improving the existing precipitation forecast quality and prediction accuracy, which has become one of the research hotspots.

The dendritic neural network has achieved great success in many fields (Wu et al., 2009, 2018; Kisi et al., 2017; Egrioglu et al., 2019; Xue et al., 2021; Achite et al., 2022). The diverse kinds of synaptic plasticity and non-linearity mechanisms enable synapses to take a valuable part in calculation (Gao et al., 2018). Synaptic non-linearity is implemented in a dendritic structure to effectively solve linearly inseparable problems, and this model has been applied to a variety of complex continuous functions (Zhou et al., 2016; Chen et al., 2017; Ji et al., 2019, 2021). Among the various types of soft computing approaches, the artificial neural networks (ANNs) models have satisfactorily been applied to non-linear hydrologic simulations such as rainfall (Acharya et al., 2014), evapotranspiration (Kisi et al., 2015), and river flow (Zounemat-Kermani et al., 2013). In recent years, related research has also been carried out at home and abroad. For example, Huating Xu (Xu et al., 2018) mentioned that the global environmental multiscale (GEM) model is widely used as a high-resolution medium-term prediction model for precipitation forecasting in various parts of Canada. With the continuous deepening of regional climate simulation research, the new generation of regional climate models, Climate-Weather Research and Forecasting (CWRF), has begun to be widely used because of its excellent performance (Guanzhou and XinZhong, 2017). For example, Xiaoyun et al. (2019) used the CWRF regional model to propose a cumulative probability transformation deviation correction method for extreme precipitation to test and evaluate its applicability to extreme precipitation correction. For example, Zhang and Zhi (2018) proposed using the frequency matching method (FMM) to calibrate the large-scale precipitation forecast data obtained from the Public Meteorological Service Center of the China Meteorological Administration (CMA). The results show that FMM calibration can significantly improve the forecasting skills of large-scale precipitation forecasts. For example, Jixue et al. (2016) used 18 meteorological elements, such as temperature, humidity, pressure, and wind field, in high-altitude weather observation data to train a three-layer BP neural network model. The experimental results show that the ANN has good application prospects in short-term precipitation forecasting. For example, Yuting et al. (2020) trained the ANN model with the annual precipitation of five stations in the western area of Taihu Lake. The results showed that the fitting and prediction accuracy and stability of the ANN model based on component analysis were higher than those of the original ANN and linear autoregression models and the other 4 types of neural networks. For example, Liu and Wang (2021) proposed a white-box dendrite net (DD) with a logical operational relationship, while the ANN network is a black-box network that does not consider the fuzzy non-linear mapping of the logical operation. DD has better generalization. Li et al. (2020) used the integrated network model of ANN and DD, and through digital recognition tasks, experiments proved the potential of artificial DDs to improve overall performance. Liangmin et al. (2016) proposed an objective clustering method based on nearest neighbor propagation to divide the climate of summer rainfall in China while using factors such as sea temperature and sea level pressure to establish a least-squares regression method to predict precipitation. Moreover, Chow and Cho (1997) described the development of new approaches to rainfall forecasting using ANN (Egrioglu et al., 2019).

In this article, the artificial DD is used to correct the CWRF simulation of summer precipitation so as to improve the accuracy of the CWRF prediction of precipitation. The data mining correlation algorithm is used for the correction of precipitation forecast results, and a precipitation correction scheme based on NN is proposed, which is used for the correction of precipitation in the summer flood season in China.

DATA AND METHODS

Data Sources

The data used were obtained from the historical return results of the CWRF regional climate model (Climate Extension of WRF) (30 km resolution) in the summer of China (June–August) from 1991 to 2020 in the National Climate Center. Among them, 7 physical configuration combinations (i.e., case 01, 02, 06, 15, 16, 23, and 28) are selected from the four starting times (i.e., 00, 06, 12, and 18) on March 2 each year, a total of 28 samples; at the same time, the actual observation data (OBS) data from the ground station during the summer of 1991–2019 is selected as the target. The main meteorological elements included are precipitation, wind speed (10 m), wind volume (10 m), relative humidity (2 m), temperature (2 m), 500 altitude field, sea level pressure, whole layer water vapor, and vertical speed.

Method Introduction

Artificial Neural Network Algorithm Model

Artificial neural network is a system with non-linear and adaptive information processing capabilities composed of a large number of neurons connected by different weights. ANN is a model that simulates the biological nervous system, which can approximate any non-linear function. There are three kinds of neurons,

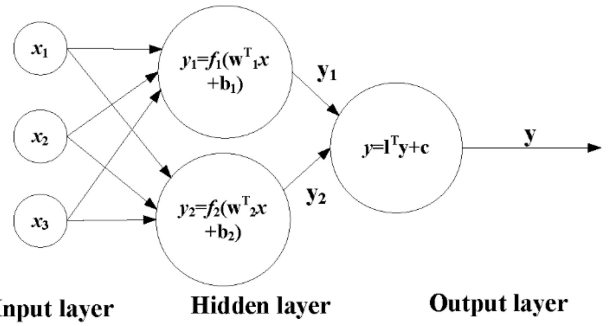


FIGURE 1 | Flowchart of artificial neural network model.

namely, (a) output neurons, those that send data out of the network; (b) input neurons, which receive external data; and (c) hidden neurons, whose signals remain in the ANN and join the input layer neurons to the neurons of the output layer (Samani et al., 2007; Zounemat-Kermani, 2012). Traditional precipitation forecasting and correction methods need to be analyzed by understanding precipitation principles and related influencing factors, while the ANN to achieve interannual precipitation forecasting does not need to be clear about the precipitation mechanism, and a model can be learned by learning precipitation and related element data to forecast future precipitation.

Artificial Neural Network Model

$$Where\ \boldsymbol{X} = [\boldsymbol{x_1}, \boldsymbol{x_2}, ..., \boldsymbol{x_n}],\ \boldsymbol{W} = \begin{bmatrix} \boldsymbol{w_{i1}} \\ \mathbf{w_{i2}} \\ \cdot \\ \boldsymbol{w_{in}} \end{bmatrix},\ \boldsymbol{b} = \begin{bmatrix} \boldsymbol{b_{i1}} \\ \boldsymbol{b_{i2}} \\ \cdot \\ \mathbf{b}_{in} \end{bmatrix}$$

The ANN algorithm model includes two parts, namely, forward propagation of information and back propagation of error. In **Figure 1**, $\mathbf{X_1} \sim \mathbf{X_2}$ are the input characteristic signals passed in from the neuron. $\mathbf{W_{i1}} \sim \mathbf{W}_{in}$ are the weights corresponding to the incoming signals of different neurons. $\mathbf{b_{i1}} \sim \mathbf{b}_{in}$ represent a bias. The setting of bias is to achieve accurate output and is an important parameter in the model. Different neurons are combined into the final input signal through different weight matrices. In **Figure 1**, $\mathbf{f}(*)$, $\mathbf{f}$ are called activation functions. The activation function mainly acts on the linear connection, and the non-linear function is added to the model, which can well realize the learning of non-linear problems. In the figure, $\mathbf{y}$ is the final output of the neuron.

Error Back-Propagation and Improvement

Back-propagation (BP) neural network is a process of continuous repetition when training the network, by collecting the errors generated by the system, returning these errors to the output value, and then using these errors to adjust the weight of the neurons so that the loss of the model propagates along the direction of the negative gradient. The parameters that affect the performance of the BP neural network mainly include the number of hidden layer nodes, the choice of activation function, and the choice of the learning rate. This article is based on the number of hidden layer nodes of the neural network to improve. According to the empirical formula:

$$\sqrt{N} + \boldsymbol{X} \tag{1}$$

Among them, $\mathbf{N}$ represents the number of sample features, and the value range of $\mathbf{X}$ is 1–10. The number of hidden layer nodes is determined step by step, and the prediction performance of each model is obtained by comparing different numbers of nodes, and the number of nodes with the best effect is selected as the number of hidden layer neurons. When determining the number of hidden layer nodes, the following conditions must be met: First, the number of hidden layer nodes must be less than $\mathbf{N-1}$, that is, less than the number of input features. Otherwise, the system error of the network model is independent of the characteristics of the training sample and tends to be zero, i.e., the constructed network model has no generalization ability. Second, the number of training samples must be more than the connection weight of the network model, generally 2–10 times. Otherwise, the sample must be divided into several parts and the method of "training in turn" can be used to obtain a reliable neural network model.

Artificial Dendrite Net Algorithm Model

Dendrite Net Algorithm Model

The main concept of the DD model is that if the output logical expression contains the logical relationship of the corresponding class between the inputs (and\or\not), the algorithm can identify the class after learning. The white box machine learning algorithm DD shows excellent system recognition performance for the black box system. The DD has white box properties, controllable accuracy, better generalization ability, and lower computational complexity. Not only can DD be used in general engineering but as a module of deep learning, it also has broad development potential. The expression of the DD module is as follows:

$$\boldsymbol{A^1} = \boldsymbol{W^{l,l-1}A^{l-1}} \circ \boldsymbol{X} \tag{2}$$

Among them, $\mathbf{A^{l-1}}$ and $\mathbf{A^l}$ represent the input and output of the model, $\mathbf{X}$ represents the input of DD, $\mathbf{W^{l,l-1}}$ is the weight matrix from module $\mathbf{l-1}$ to module $\mathbf{l}$, and $\circ$ is the multiplication of corresponding elements, sometimes called Hadamard product.

Artificial Dendrite Net Algorithm Model

The artificial dendritic neural network model is shown in **Figure 2**. The ANN algorithm model adds a DD module on the basis of ANN. The connection of different DD modules enhances the processing ability of neurons carrying information. The number of modules can effectively adjust the logic expression ability of DD and avoid excessive simulation. Together, it is easy to obtain models with outstanding generalization capabilities. By analyzing the existing ANN network model, it does not consider that the logical operation is a fuzzy non-linear mapping relationship, and the DD considers that the logical operation can converge to the global optimum with a high probability, so the integration of DD and ANN is adopted. The network model of the ANN adds a dendritic module to the hidden layer of the ANN to improve the generalization ability of the ANN algorithm model.

Evaluation Index

In this article, three indicators such as mean square error (MSE), temporal correlation coefficient (TCC), and spatial anomaly correlation coefficient (ACC), commonly used in meteorological services are used to evaluate the effect of machine learning on the correction of CWRF summer precipitation. The definitions are, respectively, as follows:

(1) The MSE is often used as an indicator to evaluate the prediction results of a machine learning model. The formula is as follows:

$$\mathbf{MSE} = \frac{1}{N}\sum_{i=1}^{N}\left(pre_i - obs_i\right)^2 \tag{3}$$

(2) The time correlation coefficient (TCC) can better represent the model's ability to predict the abnormality of each grid point in a statistical sense and obtain a complete spatial distribution of correlation techniques. When calculating TCC, the mean square deviation and covariance of each grid point are required. The formula is as follows:

$$TCC = \frac{\sum_{i=1}^{N}\left(pre_i - \overline{pre_i}\right)\left(obs_i - \overline{obs_i}\right)}{\sqrt{\sum_{i=1}^{N}\left(pre_i - \overline{pre_i}\right)^2 \sum_{i=1}^{N}\left(obs_i - \overline{obs_i}\right)^2}} \tag{4}$$

In the formula, pre_i and $\overline{pre_i}$ are the model return value or precipitation data model forecast value of **i** sample point and its multiyear average value; obs_i and $\overline{obs_i}$ are the actual observation values of the precipitation data of sample point **i**; **N** is the total number of grid points actually participating in the evaluation.

(3) The ACC mainly reflects the degree of similarity between the forecasted value and the actual value of the space type and can also be called the spatial similarity coefficient. The spatial similarity coefficient can be calculated for each forecast field. The formula for calculating ACC is as follows:

$$ACC = \frac{\sum_{i=1}^{N}\left(\Delta R_f - \overline{\Delta R_f}\right)\left(\Delta R_o - \overline{\Delta R_o}\right)}{\sqrt{\sum_{i=1}^{N}\left(\Delta R_f - \overline{\Delta R_f}\right)^2 \sum_{i=1}^{N}\left(\Delta R_o - \overline{\Delta R_o}\right)^2}} \tag{5}$$

In the formula, $\mathbf{\Delta R_f}$ and $\overline{\mathbf{\Delta R_f}}$ are the forecast value and multiyear average value of precipitation anomaly percentage ((actual measured value-historical average value of the same period)/historical average value of the same period); $\mathbf{\Delta R_o}$ and $\overline{\mathbf{\Delta R_o}}$ are the corresponding observation values; **N** is the total number of stations actually participating in the evaluation.

The formula for the abnormal comprehensive (Ps) test score, which is a commonly used predictive scoring index in business, is:

$$P_S = \frac{a*N0 + b*N1 + c*N2}{(N - N0) + a*N0 + b*N1 + c*N2 + M} * 100 \tag{6}$$

The scoring steps are as follows:

1. Determine whether the forecasted trend is correct from station to station and calculate the total number of stations with correct trend prediction **N0**;
2. Determine whether the first-level anomaly prediction is correct from station to station and calculate the total number of stations **N1** with the correct first-level anomaly prediction;
3. Determine whether the second-level anomaly prediction is correct from station by station and count the total number of stations **N2** with the correct second-level anomaly prediction;
4. The number of stations where the percentage of precipitation anomaly is $\geq$ 100% or equal to $-$100%, and the temperature anomaly is $\geq$ 3°C or $\leq$ $-$3°C (referred to as missed stations, denoted as **M**) without a second-level anomaly forecast;
5. Count the number of stations **N** that actually participated in the assessment, that is, the number of stations that are required to participate in the assessment minus the number of stations that are not in the live test;

a, **b**, and **c** are the weight coefficients of the climate trend item, the first-level anomaly item, and the second-level anomaly item, respectively. This algorithm takes **a** = **2**, **b** = **2**, and **c** = **2**, respectively.

PRECIPITATION CORRECTION MODEL CONSTRUCTION

National Climate Division

In this study, the precipitation prediction results of the CWRF model in the flood season (June–August) in China (8.37°N–58.75°N, 58.40°E–161.60°E) were selected as the target of forecast correction. The geographical diversity of climate is obvious, which makes it difficult for a single model to represent the climate characteristics of the entire China, leading to certain difficulties in climate forecasting. To realize the correction of the national regional precipitation forecast data, regional modeling forecasts are carried out for the whole country. That is, according to the climatic characteristics of different regions, they establish model algorithms suitable for their respective climatic characteristics, promote the high generalization ability of the model, and improve the accuracy of precipitation forecasting. According to the climate characteristics of different regions, China is divided into eight regions (Wang and Yang, 2017). The specific results are shown in **Table 1**.

In the precipitation data file, the data size is 231 × 171, that is, 231 × 171 grid points, including the China region. Now, the regions are divided according to **Figure 3** and **Table 1**.

Feature Selection and Data Organization

Feature Selection

Pearson correlation coefficient (PCC) and random forest algorithm for feature selection comparison.

Pearson correlation coefficient, also known as Pearson product-moment correlation coefficient (PPMCC or PCCs), is

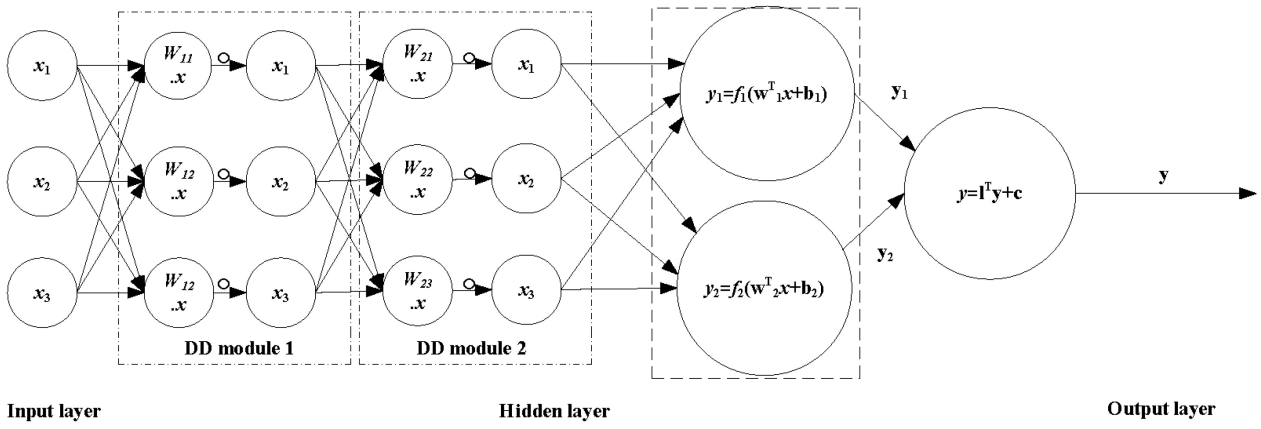


FIGURE 2 | Flowchart of artificial dendritic neural network model.

used to measure the correlation between two variables **X** and **Y** (linear correlation), and its value is between -1 and 1.

The PCC between two variables is defined as the quotient of the covariance and standard deviation between the two variables:

$$\rho_{X,Y} = \frac{cov(X, Y)}{\sigma_X \sigma_Y} = \frac{E[(X-\mu_X)(Y-\mu_Y)]}{\sigma_X \sigma_Y} \quad (7)$$

The above formula defines the overall correlation coefficient, and Greek lowercase letters are commonly used as representative symbols. Estimate the covariance and standard deviation of the sample to get the PCC. Commonly used English lowercase letter *r* stands for:

$$r = \frac{\sum_{i=1}^{n} (X_i-\overline{X})(Y_i-\overline{Y})}{\sqrt{\sum_{i=1}^{n} (X_i-\overline{X})^2}\sqrt{\sum_{i=1}^{n} (Y_i-\overline{Y})^2}} \quad (8)$$

r can also be estimated from the mean value of the standard scores of the sample points ($\mathbf{X_i}$, $\mathbf{Y_i}$), and an expression equivalent to the above equation can be obtained:

$$r = \frac{1}{n-1} \sum_{i=1}^{n} \left(\frac{X_i-\overline{X}}{\sigma_X}\right)\left(\frac{Y_i-\overline{Y}}{\sigma_Y}\right) \quad (9)$$

Among them, $\frac{X_i-\overline{X}}{\sigma_X}$, $\overline{X}$, and σ_X are the standard score, sample mean, and sample standard deviation of $\mathbf{X_i}$ sample, respectively.

TABLE 1 | Eight area names.

Abbreviations	Full name
NWCH	Northwest China
TP	Tibetan Plateau
BBYR	Big Bend of Yellow River
SWCH	Southwest China
NECH	Northeast China
NCH	North China
YHRB	Yangtze–Huaihe River Basin
SCH	South China

1. Use PCC to select features for precipitation data.
2. Use random forest model to select features.

After experiments, for the random forest model, using the attribute column obtained through the PCC in step 1 for training, the score is 0.97; while using the features selected by the random forest for training, the score is 0.98. It can be seen that the use of random forest for feature selection still has a certain effect on improving the ability of the model on this dataset.

Later, we picked the characteristics that are most important to us: the precipitation and the historical return of the CWRF. These are cumulative wind volume v component, cumulative wind speed, cumulative temperature 2 m, and precipitation correlation of the historical return.

Data Organization

First of all, considering the influence of the age average on forecasting, anomalies are commonly used in meteorological forecasting operations, so this article also chooses anomalies to preprocess the data. Taking precipitation as an example, comparing the rainfall forecasted by the model with the average rainfall over the years, the forecast is the value of the forecast

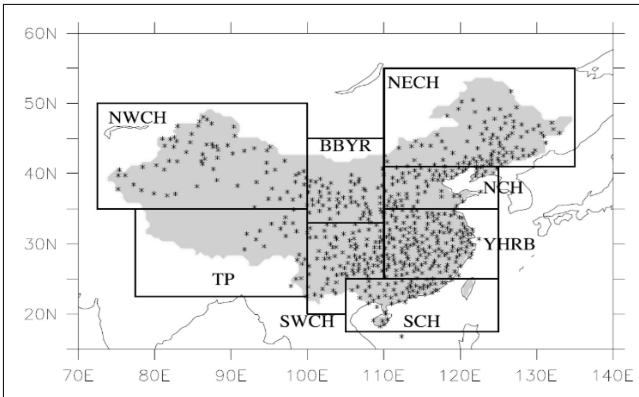


FIGURE 3 | China's terrestrial climate regional division map.

rainfall minus the average rainfall over the same period. It is generally used in medium- and long-term forecasts and can be used as a reference for flood control and drought resistance. The historical year-month (year) precipitation distance is equal to the difference between the historical year-month (year) precipitation and the cumulative year-month (year) average precipitation.

$$Dpt = Pre - AvgPre \tag{10}$$

Among them, Pre represents the precipitation of a certain location in a certain month in a certain year, AvgPre represents the average precipitation of a certain location in the historical years that have been recorded at that point, and Dpt represents the precipitation of a certain location in a certain year from the average value in a certain month. If Dpt $<$ **0**, it is a positive anomaly, and the annual precipitation at the location is greater than the cumulative annual average precipitation; if Dpt $>$ **0**, it is a negative anomaly, and the annual precipitation at the location is less than the cumulative annual average precipitation.

Then, the summer precipitation data are gridded data. Each grid point has only the CWRF return data and actual observation data at a specific time and a specific location. Based on the similarity characteristics of the climate in the neighboring areas, the target point is divided into small areas of **M**∗**M**, and then each grid point has **M**∗**M** feature data; (1) the precipitation is corrected by the single-element integration method: the value of M is 3, which is a small area of 3 × 3 around the grid point. Second, taking into account the influence of the precipitation months in summer (6–8), the average precipitation anomaly of the past 5 months (4–8) is used as the input feature of the model to organize the data. That is, the April–August precipitation anomaly and the average precipitation of the 3 × 3 grid points around the current year of the CWRF model precipitation are used as the characteristics of the model training input, and the output is the precipitation anomaly from April to August of the current year; (2) using the multifactor integration method to correct the precipitation: select the historically reported precipitation of CWRF, the cumulative wind volume u component, the cumulative wind volume v component, the cumulative wind speed, and the cumulative temperature 2 m. The data are preprocessed and converted into monthly anomaly data. Five meteorological elements, input according to 3 × 3 grid points are selected; the input data from June to August (or a single month) are organized, and the corresponding 1-month precipitation anomaly is delivered. Data according to the above different methods are organized, and a data format suitable for network model training is constructed. Finally, based on the interdecadal influence in the meteorological field, this article uses the data of nearly **N** years when training based on the artificial DD model and then predicts the summer precipitation in **N1** years. The range of **N** selected in this article is 3–10. The analysis of the experimental comparison results shows that **N** is set to 5, that is, the training algorithm model using the data of the past 5 years is better than the correction results of other years.

Through the above-mentioned modeling method, the forecast factors are optimized at the same time for the 28 sample forecast data of 7 different physical parameter configurations and 4 start times. One or more forecast factors are used to predict the precipitation, and the 28 sample forecast results use the weighted average method to integrate as the final output. First, the organized data are divided as follows: 70% of the data as the training set, 20% of the data as the validation set, and the rest as the test set. At the same time, the MSE is selected as the error loss function for network model training, based on the artificial DD model summer precipitation. The flowchart is shown in **Figure 4**.

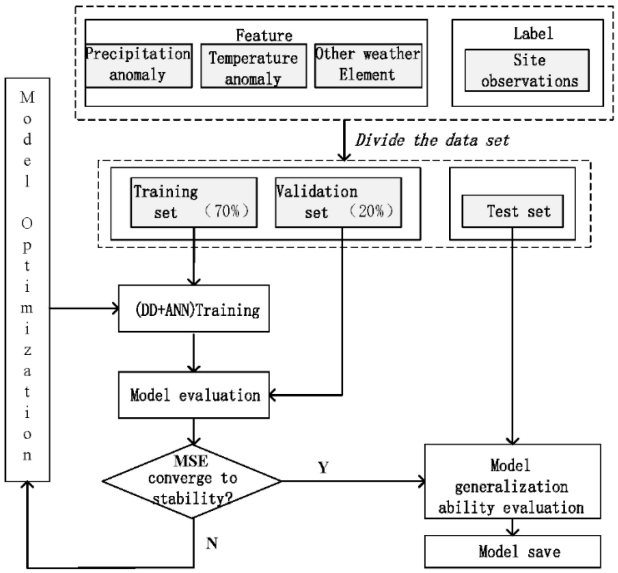


FIGURE 4 | Flowchart of summer precipitation correction algorithm based on artificial dendrite net.

ANALYSIS OF PRECIPITATION CORRECTION RESULTS

Regional Precipitation Forecast Revision

(1) Comparison of TCC before and after the regional correction in China.

The eight sub-areas are divided to calculate the TCC of each grid point, and the correction results of the precipitation forecast of the flood season model are tested. **Figures 5A,B** shows the TCC comparison test of precipitation and actual observed precipitation (OBS) in the eight sub-regions during the flood season from 1996 to 2019. It can be seen from **Figures 5A,B** that after the machine learning method is revised, the positive value range of the TCC of precipitation prediction in different regions has been improved to varying degrees, especially in the eastern key areas of the rain belt during the flood season in China [South China (SCH), Yangtze-Huaihe River Basin (YHRB), North China (NCH), Northeast China (NECH)] TCC correction effect is more obvious. The YHRB, the NECH, and the Southwest Region (SWCH) have significantly improved TCC compared with other regions. After the revision, the regions that passed the 90% significance test significantly increased.

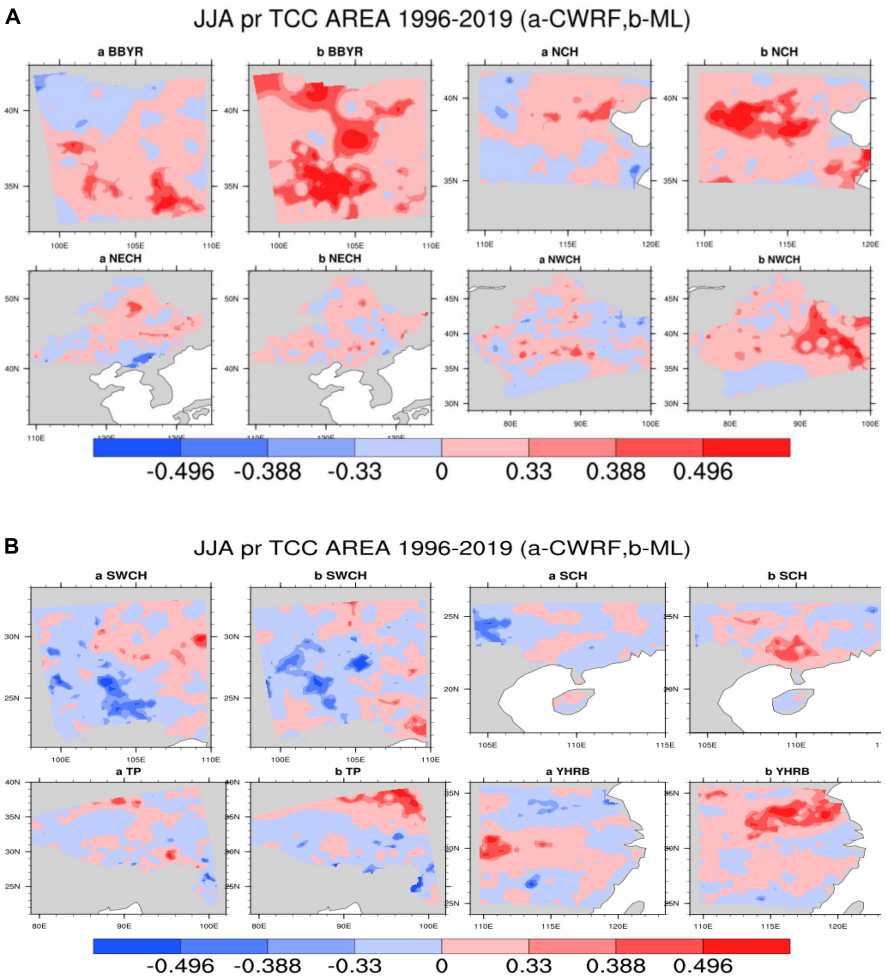


FIGURE 5 | (A) Comparison of regional precipitation time correlation coefficients [Climate-Weather Research and Forecasting (CWRF) vs. Machine Learning (ML)], the color code numbers 0.33, 0.388, and 0.496, respectively, represent that the correlation coefficients have passed the significance test of 90%, 95%, and 99%, respectively. **(B)** Comparison of regional precipitation time correlation coefficients (CWRF vs. ML); the color code numbers are the same as panel **(A)**.

It can be seen from **Figure 6** that after the correction of the machine learning method, the mean value of the TCC of the precipitation prediction in different regions has been improved to different degrees. The range of the positive value after the correction by the improved ANN algorithm has been improved more obviously, especially in the rainy season in China. The effect of TCC correction in key eastern regions of the belt [Big Bend of Yellow River (BBYR), Northwest China (NWCH), SCH, YHRB, NCH, and NECH] is more obvious. The Jianghuai River Basin (YHRB), BBYR, NCH, and NWCH have significantly improved TCC compared with other regions. From the regional average, the difference between the results of experiments before and after the correction has more than 0.2.

(2) Comparison of spatial correlation coefficients (i.e., ACC) before and after the revision of China's regional precipitation forecasts.

Calculate the spatial correlation coefficients between the precipitation during the flood season from 1996 to 2019 and the actual observed precipitation for the eight sub-regions divided, and the ACC values before and after the correction are shown in **Figure 7**. It can be seen from **Figure 7** that the spatial correlation coefficient of the corrected results of ML in the 8 regions is larger than the original prediction of CWRF. The spatial correlation coefficient of b is relatively close. After the machine learning method is revised, the spatial correlation coefficients of precipitation forecasts in different regions have been improved to varying degrees. The algorithm has certain predictive performance.

It can be seen from **Figure 8** that after the correction of the machine learning method, the mean values of the spatial correlation coefficients of precipitation forecasts in different regions have improved to varying degrees. Among them, BBYR, NCH, NWCH, SCH, SWCH, and YHRB have improved significantly, especially during the flood season in China. The ACC correction effect in the key eastern areas of the rain belt (i.e., SCH, YHRB, NCH, NECH, and SWCH) is more obvious. The Jianghuai River Basin (YHRB), the NECH, and the Southwest Region (SWCH) have significantly improved ACC

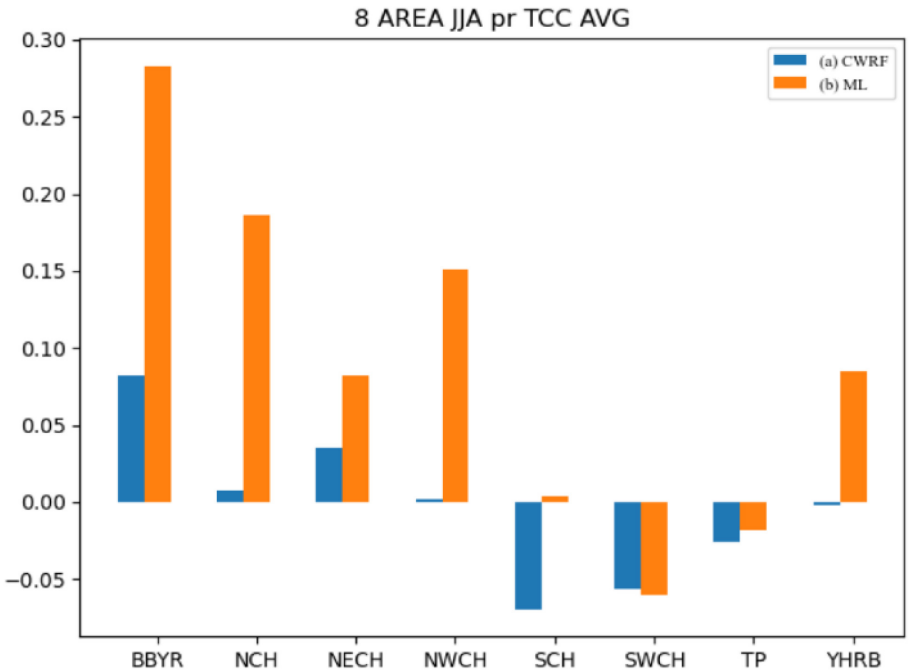


FIGURE 6 | Comparison of mean values of regional precipitation time correlation coefficients (CWRF vs. ML). (a) CWRF prediction and (b) machine learning correction.

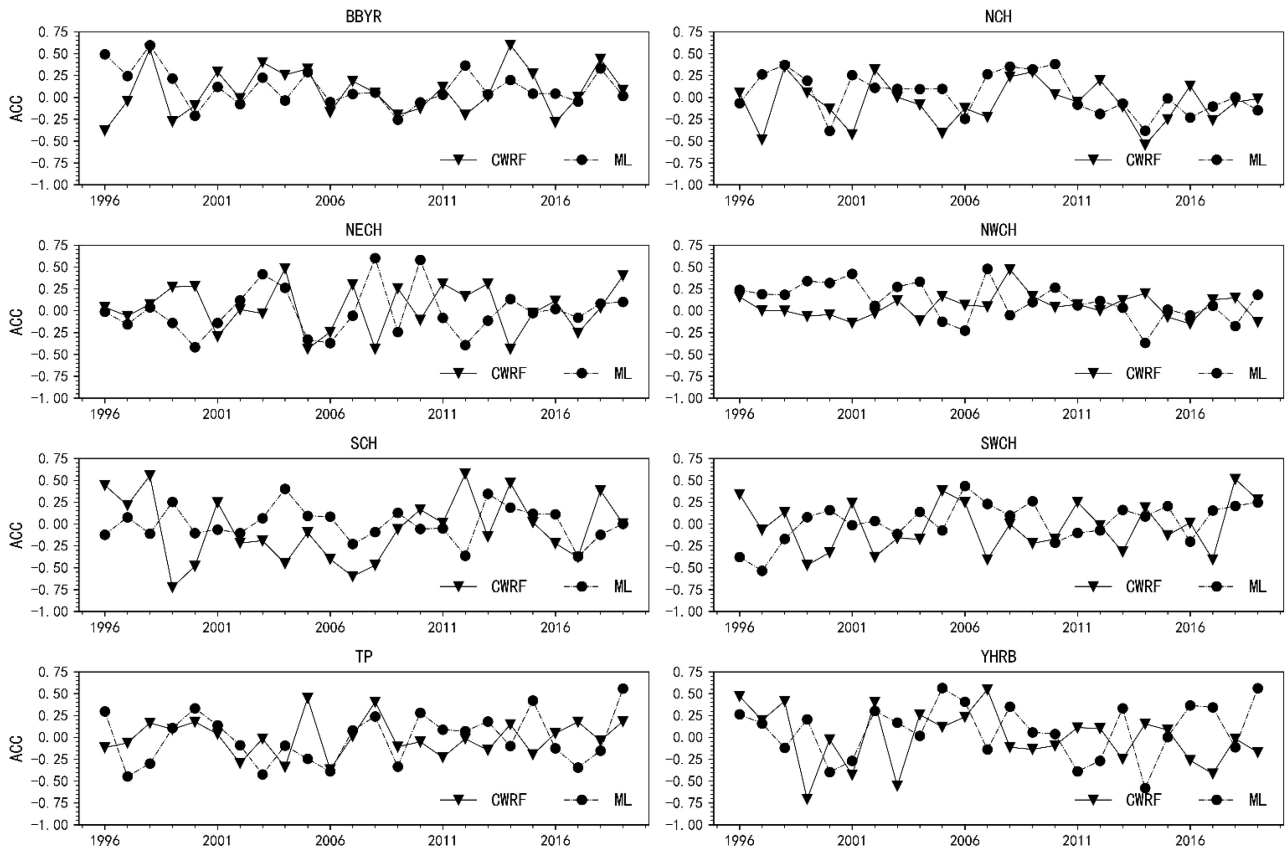


FIGURE 7 | Comparison of anomaly correlation coefficient (ACC) before and after precipitation correction in 8 sub-regions.

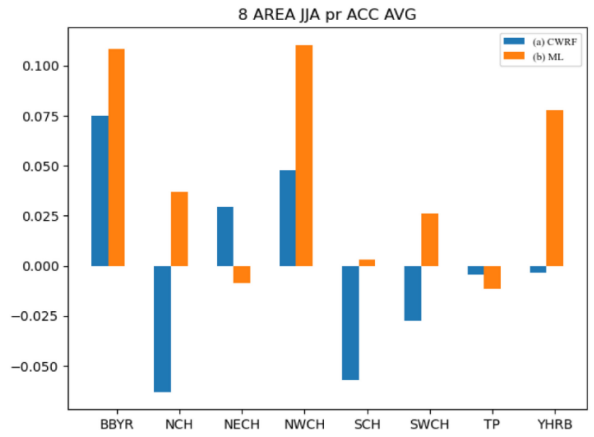


FIGURE 8 | Comparison of mean values of spatial correlation coefficients of regional precipitation (CWRF vs. ML). (a) CWRF prediction and (b) machine learning correction.

compared with other regions. From the regional average, the difference in results before and after the correction has risen more than 0.1 on average.

Evaluation of the Effect of the National Precipitation Forecast Revision

(1) Comparison of the MSE before and after the correction of the national summer precipitation forecast (puzzle).

Figure 9 contains the comparison of two calculation results, namely, (1) MSE change (CWRF) between simulated precipitation (case) and actual observed precipitation (OBS) in CWRF model and (2) use artificial network to correct precipitation and actual observation (OBS) between MSE change (ML) under the corresponding parameters; from **Figure 9**, it can be seen that there are fluctuations between the different CWRF model data and the MSE of the actual observation data, indicating that different physical parameter configurations have different degrees of error in the simulated flood season precipitation; after correction, the rainy season precipitation of the next 7 cases all increased to varying degrees. The 7 case results of CWRF simulated precipitation, respectively, calculate the MSE, and the average MSE is 9.45, and the corrected average MSE is 7.03, a decrease of 2.42 (equivalent to a 26% decrease in MSE). It can be seen from the figure that the MSE of the CWRF model rainy season precipitation data fluctuates greatly, and the corrected 7 case results have stabilized, indicating that the artificial network model has improved the accuracy of the model forecast data to a certain extent. The forecasting performance of CWRF has to be improved.

(2) Comparison of the TCC before and after the correction of the national summer precipitation forecast (puzzle).

Figure 10 shows the TCC comparison before and after the correction of CWRF forecasts for 7 cases and 4 time-time sets (28 samples in total). It can be seen from the figure that there are fewer positive correlation areas for TCC before the correction, especially in eastern China. There is a large area of negative value in the key area (the Yangtze River Basin to SCH). After the correction of the artificial network model, the TCC of the key precipitation area in eastern China shows a large positive correlation, especially in the south of the Yangtze River and the northeast. After correction by machine learning, the number of positive TCC regions and significant regions increased significantly across the country, especially in the Yangtze River Basin and Northeast China, indicating that the ANN-DD algorithm model can achieve certain corrections to the national regional precipitation forecast data during the flood season. Forecasting skills have been significantly improved.

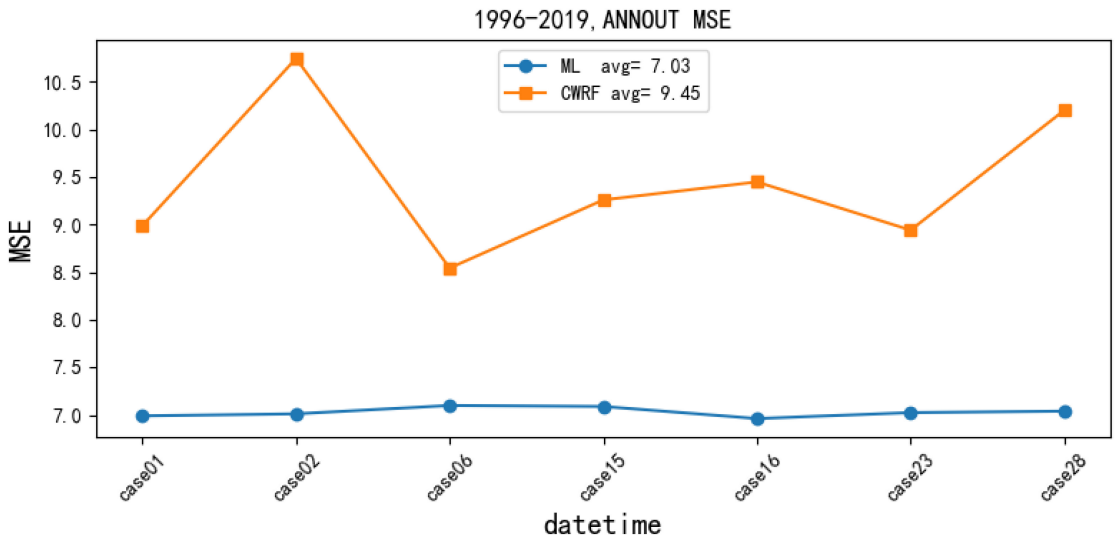


FIGURE 9 | Mean square error (MSE) change graph of precipitation data correction results for 7 cases (i.e., 01, 02, 06, 15, 16, 23, and 28) of CWRF based on the artificial DD model.

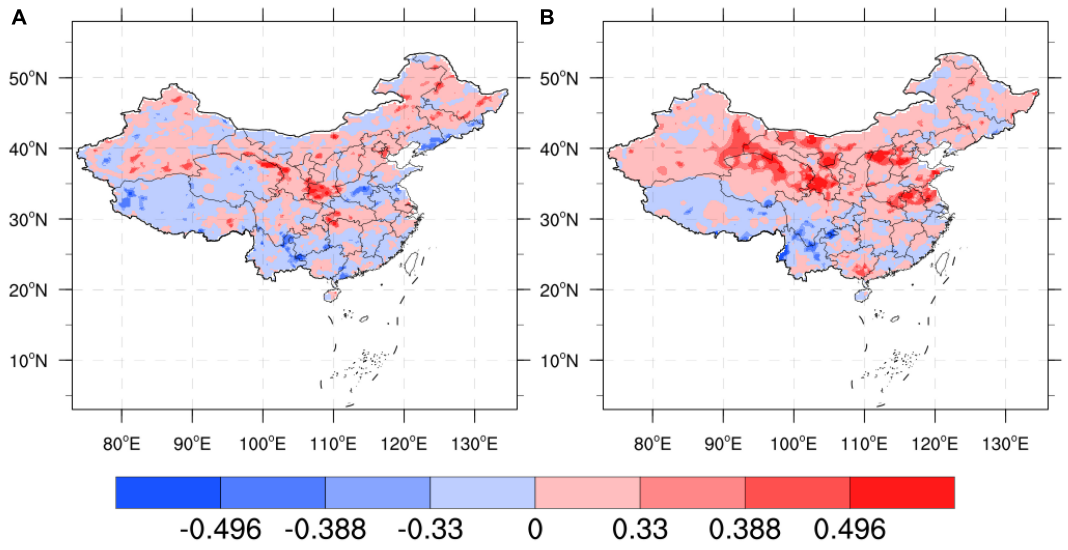


FIGURE 10 | Comparison of TCC before and after CWRF prediction correction for 7 case sets, **(A)** CWRF prediction and **(B)** machine learning correction, the color code number is the same as **Figure 4**.

(3) Comparison of the spatial correlation coefficient (ACC) before and after the correction of the national summer precipitation forecast (puzzle).

Figure 11 shows the ACC comparison before and after the correction of CWRF forecasts for 7 cases and 4 time sets (28 samples in total). It can be seen from the figure that the average ACC before correction is −0.01, and the ACC value in most years is low at 0, there is a negative correlation. After correction by machine learning, the ACC value has increased significantly. The correction result of the ML (curve b in the figure) scheme is about 0.1 higher than before the correction, and the ACC value in most years is greater than 0. At the same time, the predicted result after correction can be seen in the figure. The stability has improved, and the fluctuations before correction are large, showing a good correction effect. It shows that the ANN algorithm model can make certain corrections to the national regional rainy season precipitation forecast data, and the forecasting skills have improved.

(4) Comparison of comprehensive trend anomaly scores (Ps) before and after correction of the national summer precipitation forecast (puzzle).

Consider case 28 (**Figure 12**) and 7 case set (**Figure 13**) as examples to illustrate the improvement effect of the CWRF model prediction Ps of the machine learning correction method.

Figure 12 shows the comparison of the CWRF prediction and the Ps score corrected by ANN after the 4 time collections of case 28 (the 4 time precipitation predictions of case 01 are collected before calculating the Ps). It can be seen from the figure that the average value of Ps before correction is around 70.87, and after correction by machine learning, the Ps score is around 72.55, and the overall prediction skills have improved. Among them, ML (curve b in the figure) has a better overall trend abnormality score in 1996–2014 than CWRF's original prediction (curve a in the figure); the Ps score in 2015–2019 fluctuates more. This means that the simulated precipitation of case 28 is better because of machine learning.

Figure 13 shows the comparison of CWRF prediction and ANN corrected Ps after the collection of 4 starting times for 7 cases (before calculating Ps, the precipitation predictions for 4 times (28 samples in total) of 7 cases are collected) deal with. The CWRF forecast on the graph shows that the average Ps score from 1996 to 2019 is about 67.79; the average Ps score after machine learning correction is about 74.34; **Figure 13** generally shows the same correction effect as **Figure 12**, except for the Ps value of individual years' difference. This shows that the precipitation forecasting skills during the flood season have improved to a certain extent compared with before the

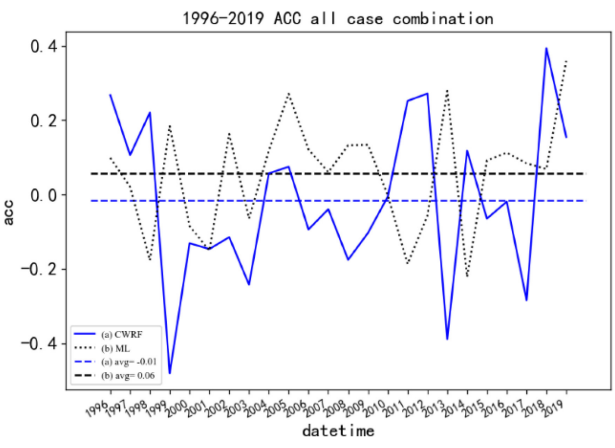


FIGURE 11 | Comparison of ACC before and after CWRF forecast correction for 7 case sets. (a) CWRF forecast, (b) machine learning correction, and (a, b) 1996–2019 average.

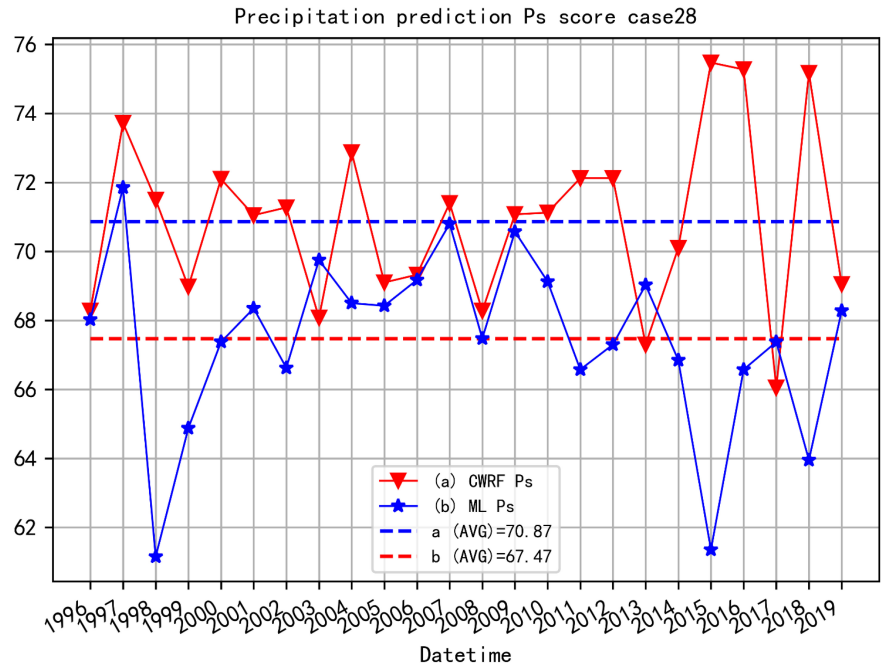


FIGURE 12 | Comparison of Ps before and after CWRF forecast correction for the 4 time-time sets of case 28. (a) CWRF forecast, (b) machine learning correction, and (a, b) mean Ps from 1996 to 2019.

correction, and the Ps score of most years after the correction has improved. The Ps score shows a stable trend with a small fluctuation range. Based on the above two calculations of Ps anomaly scores for different physical parameter configurations and integrations, it can be concluded that the prediction performance of the revised algorithm model based on machine learning for the rainy season precipitation prediction results is improved compared with the climate model simulation, which can achieve a certain degree of improvement in the rainy season precipitation forecast skills.

In summary, it can be concluded from **Figure 13** that the MSE, Ps, ACC, and TCC values before and after the CWRF prediction are improved, and the prediction performance of the summer precipitation prediction result correction algorithm model based on the dendritic neural network model can be obtained. Compared with the CWRF climate model simulation, it is improved and can realize the correction of summer precipitation forecast data to a certain extent.

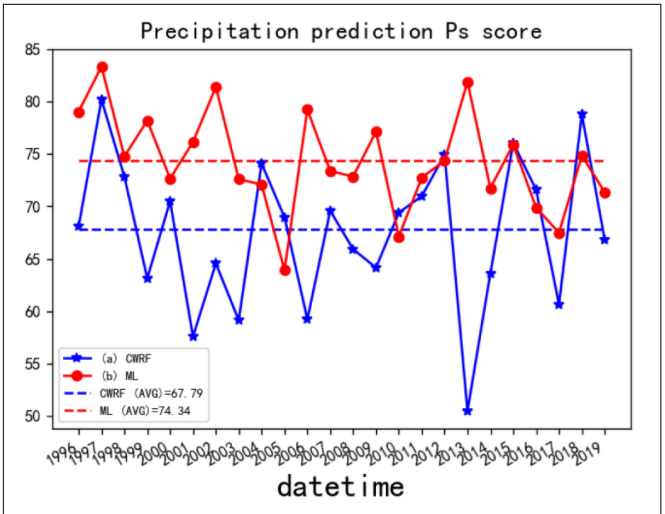


FIGURE 13 | Comparison of Ps before and after CWRF forecast correction for 7 case sets. (a) CWRF forecast, (b) machine learning correction, and (a, b) mean Ps from 1996 to 2019.

CONCLUSION AND FUTURE WORK

In this article, we used the information from nearby regions and time series and latitude and longitude positions to organize the data and then divide it into eight regions according to the climate characteristics of different regions. Then, the ANN model is improved, and the dendritic module is introduced to improve the generalization ability of the model. When constructing the network model in this article, the number of hidden layers is 2, and the number of neurons in each layer is 9. The back propagation algorithm is used for model training. Through cross-validation experiments on the model, it is found that the activation function of the hidden layer is adopted by the RELU function. The optimization algorithm adopts the Adam algorithm model to predict better. The CWRF regional climate model simulated summer precipitation was corrected by optimizing the relevant meteorological elements such as precipitation and temperature and analyzed with related

evaluation indicators such as MSE, TCC, and ACC. It was found that all three indicators were improved, indicating that the artificial dendritic network model is effective for CWRF. The accuracy of model forecast data has been improved, which can improve the forecasting performance of CWRF to a certain extent. The experimental results are obtained using the cross-validation method, which can objectively evaluate the generalization performance of the model. Experiments show that good results can be achieved, which makes this method a good choice for the meteorological field.

This article mainly discusses the problem of precipitation correction in flood season. In theory, this idea and method can be used in other related meteorological forecasting fields and can be used as future work. At present, the model only considers the output of the climate model, and many other related elements are not used for modeling, so the advantages of big data are not fully utilized. In the future, more meteorological elements can be introduced to overcome this limitation. This article cannot remove the impact of extreme weather precipitation data when using precipitation data, and only the forecast output of the meteorological model is used as the modeling object of the algorithm in this article. Taking into account the complexity of meteorological problems. Furthermore, we hope to reduce the impact of extreme weather by layered modeling of precipitation. Other data, such as relevant observation data, can be added as model inputs to check whether the accuracy of the model can be improved.

DATA AVAILABILITY STATEMENT

The data analyzed in this study is subject to the following licenses/restrictions: It is the business data provided by the relevant departments of the China Meteorological Administration, non-public data, and needs to be applied to the relevant departments of the China Meteorological Administration. Requests to access these datasets should be directed to https://data.cma.cn/.

AUTHOR CONTRIBUTIONS

TL and CQ responsible for designing and implementing algorithms. LW and JC responsible for organize experiment. YR responsible for the analysis of experimental results and algorithm optimization. JC responsible for language polish. All authors contributed to the article and approved the submitted version.

FUNDING

This study received funding from the National Natural Science Foundation of China under Grant Nos. 61772280 and 62072249. The funder was not involved in the study design, collection, analysis, interpretation of data, the writing of this article, or the decision to submit it for publication.

REFERENCES

Acharya, N., Shrivastava, N. A., Panigrahi, B. K., and Mohanty, U. C. (2014). Development of an artificial neural network based multi-model ensemble to estimate the northeast monsoon rainfall over south peninsular India: an application of extreme learning machine. *Clim. Dynam.* 43, 1303–1310. doi: 10.1007/s00382-013-1942-2

Achite, M., Banadkooki, F. B., Ehteram, M., Bouharira, A., Ahmed, A. N., and Elshafie, A. (2022). Exploring Bayesian model averaging with multiple ANNs for meteorological drought forecasts. *Stoch. Environ. Res. Risk Assess.* 36, 1–26.

Bannister, D., Orr, A., Jain, S. K., Holman, I. P., Momblanch, A., Phillips, T., et al. (2019). Bias correction of high-resolution regional climate model precipitation output gives the best estimates of precipitation in Himalayan catchments. *J. Geophys. Res.* 124, 14220–14239. doi: 10.1029/2019jd030804

Chen, W., Sun, J., Gao, S., Cheng, J. J., Wang, J., and Todo, Y. (2017). Using a single dendritic neuron to forecast tourist arrivals to Japan. *IEICE Trans. Inf. Syst.* 100, 190–202. doi: 10.1587/transinf.2016edp7152

Chow, T. W. S., and Cho, S. Y. (1997). Development of a recurrent Sigma-Pi neural network rainfall forecasting system in Hong Kong. *Neural Comput. Appl.* 5, 66–75. doi: 10.1007/bf01501172

Egrioglu, E., Yolcu, U., Bas, E., and Dalar, A. Z. (2019). Median-Pi artificial neural network for forecasting. *Neural Comput. Appl.* 31, 307–316. doi: 10.1007/s00521-017-3002-z

Gao, S., Zhou, M., Wang, Y., Cheng, J., Yachi, H., and Wang, J. (2018). Dendritic neuron model with effective learning algorithms for classification, approximation, and prediction. *IEEE Trans. Neural Netw. Learn. Syst.* 30, 601–614. doi: 10.1109/TNNLS.2018.2846646

Guanzhou, L., and XinZhong, L. (2017). Progress of the climate extension of weather research and forecast(CWRF) model application in China. *Adv. Earth Sci.* 32:781.

Ji, J., Dong, M., Lin, Q., and Tan, K. C. (2021). Forecasting wind speed time series via dendritic neural regression. *IEEE Comput. Intell. Mag.* 16, 50–66. doi: 10.1109/mci.2021.3084416

Ji, J., Song, S., Tang, Y., Gao, S., Tang, Z., and Todo, Y. (2019). Approximate logic neuron model trained by states of matter search algorithm. *Knowl. Based Syst.* 163, 120–130. doi: 10.1016/j.knosys.2018.08.020

Jixue, Z., Peng, W., Lin, Z., and Yi, W. (2016). Research on the application of artificial neural network in short-term precipitation prediction. *Technol. Wind* 41, 123–124. doi: 10.1016/b978-0-12-801559-9.00006-5

Kim, G., Cha, D. H., Lee, G., Park, C., Jin, C. S., Lee, D. K., et al. (2020). Projection of future precipitation change over South Korea by regional climate models and bias correction methods. *Theor. Appl. Climatol.* 141, 1415–1429. doi: 10.1007/s00704-020-03282-5

Kim, K. B., Bray, M., and Han, D. (2015). An improved bias correction scheme based on comparative precipitation characteristics. *Hydrol. Process.* 29, 2258–2266. doi: 10.1002/hyp.10366

Kisi, O., Alizamir, M., and Zounemat-Kermani, M. (2017). Modeling groundwater fluctuations by three different evolutionary neural network techniques using hydroclimatic data. *Nat. Hazards* 87, 367–381. doi: 10.1007/s11069-017-2767-9

Kisi, O., Sanikhani, H., Zounemat-Kermani, M., and Niazi, F. (2015). Long-term monthly evapotranspiration modeling by several data-driven methods without climatic data. *Comput. Electron. Agric.* 115, 66–77. doi: 10.1016/j.compag.2015.04.015

Li, X., Tang, J., Zhang, Q., Gao, B., Yang, J. J., Song, S., et al. (2020). Power-efficient neural network with artificial dendrites. *Nat. Nanotechnol.* 15, 776–782. doi: 10.1038/s41565-020-0722-5

Liangmin, D., Zongjian, K., Changzheng, L., Ying, X., and Lvliu, L. (2016). Summer precipitation prediction models based on the clustering regionalization in China. *Meteorol. Mon.* 42, 89–96.

Liu, G., and Wang, J. (2021). Dendrite net: a white-box module for classification, regression, and system identification. *IEEE Trans. Cybern.* 1–14.

Pan, X., Yang, D., Li, Y., Barr, A., Helgason, W., Hayashi, M., et al. (2016). Bias corrections of precipitation measurements across experimental sites in different ecoclimatic regions of western Canada. *Cryosphere* 10, 2347–2360. doi: 10.5194/tc-10-2347-2016

Ren, Z., and Li, M. (2007). Errors and correction of precipitation measurements in China. *Adv. Atmos. Sci.* 24, 449–458. doi: 10.1007/s00376-007-0449-3

Rummukainen, M. (2010). State-of-the-art with regional climate models. *Wiley Interdiscip. Rev.*1, 82–96. doi: 10.1002/wcc.8

Samani, N., Gohari-Moghadam, M., and Safavi, A. A. (2007). A simple neural network model for the determination of aquifer parameters. *J. Hydrol.* 340, 1–11. doi: 10.1016/j.jhydrol.2007.03.017

Wang, F., and Yang, S. (2017). Regional characteristics of long-term changes in total and extreme precipitations over China and their links to atmospheric–oceanic features. *Int. J. Climatol.* 37, 751–769. doi: 10.1002/joc.4737

Wu, C., Chau, K. W., and Li, Y. S. (2009). Methods to improve neural network performance in daily flows prediction. *J. Hydrol.* 372, 80–93. doi: 10.1016/j.jhydrol.2009.03.038

Wu, X., Liu, X., Li, W., and Wu, Q. (2018). Improved expressivity through dendritic neural networks. *Adv. Neural Inf. Process. Syst.* 2018:31. doi: 10.1016/j.neuroscience.2021.10.001

Xiaoyun, D., Jinhua, Y., Xinzhong, L., and Yuan, M. (2019). Error correction of CWRF model in summer extreme precipitation simulations in China. *J. Appl. Meteorol. Sci.* 30, 223–232.

Xinmin, Z., Shiwei, X., and Linghua, Y. (2019). Discussion on the function of weather forecast on agricultural disaster prevention and mitigation. *Agric. Technol. Inf.* 22, 30-33.

Xu, H., Wu, Z., Luo, L., and He, H. (2018). Verification of high-resolution medium-range precipitation forecasts from Global Environmental Multiscale Model over China during 2009–2013. *Atmosphere* 9:104. doi: 10.3390/atmos9030104

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

Yao, T., Xuejie, G., Zhenyu, H., and Ying, X. (2017). Bias correction of daily precipitation simulated by regcm4 model over china. *Chinese J. Atmos. Sci.* 1156–1166.

Yingdong, Y., Zhiyong, Y., Dayong, Q., and Hao, W. (2011). "Appropriate spatial scale analysis for the simulation of precipitation by Regional Climate Model," in *Proceedings of the 2011 International Conference on Remote Sensing, Environment and Transportation Engineering*. (Piscataway, NJ: IEEE), 2967–2970.

Yuting, Z., Dong, W., Yuankun, W., Wenpeng, W., and Deqing, M. (2020). Multi-indexes optimization of typical artificial neural networks for rainfall forecasting: a case study in west taihu lake basin. *J. China Hydrol.* 1.

Zhang, H., and Zhi, X. (2018). "Calibration of extended-range precipitation forecasting over china," in *Proceedings of the 2018 5th International Conference on Systems and Informatics (ICSAI)*. (Piscataway, NJ: IEEE), 1254–1258.

Zhou, T., Gao, S., Wang, J., Chu, C., Todo, Y., and Tang, Z. (2016). Financial time series prediction using a dendritic neuron model. *Knowl. Based Syst.* 105, 214–224. doi: 10.1016/j.knosys.2016.05.031

Zounemat-Kermani, M. (2012). Hourly predictive Levenberg–Marquardt ANN and multi linear regression models for predicting of dew point temperature. *Meteorol. Atmospheric Phys.* 117, 181–192. doi: 10.1007/s00703-012-0192-x

Zounemat-Kermani, M., Kisi, O., and Rajaee, T. (2013). Performance of radial basis and LM-feed forward artificial neural networks for predicting daily watershed runoff. *Appl. Soft Comput.* 13, 4633–4644. doi: 10.1016/j.asoc.2013.07.007

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.

The handling editor declared a shared affiliation with the authors [CQ, TL, LW, JC, and YR] at the time of review.

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: 27 June 2022
doi: 10.3389/fpls.2022.916474

Tasseled Crop Rows Detection Based on Micro-Region of Interest and Logarithmic Transformation

Zhenling Yang[1], Yang Yang[1,2], Chaorong Li[3], Yang Zhou[1], Xiaoshuang Zhang[1], Yang Yu[1] and Dan Liu[4]*

[1] School of Engineering, Anhui Agricultural University, Hefei, China, [2] Institute of Artificial Intelligence, Hefei Comprehensive Nation Science Center, Hefei, China, [3] Faculty of Artificial Intelligence and Big Data, Yibin University, Yibin, China, [4] JD AI Research, Beijing, China

OPEN ACCESS

Edited by:
Yu Xue,
Nanjing University of Information Science and Technology, China

Reviewed by:
Junfeng Chen,
Hohai University, China
Xijian Fan,
Nanjing Forestry University, China
Xiaoli He,
Sichuan University of Science and Engineering, China
Grandee Lee,
National University of Singapore, Singapore

***Correspondence:**
Yang Yang
yy2016@ahau.edu.cn

Specialty section:
This article was submitted to Intelligent Computing Research with Applications in Ecological Plant Protection, a section of the journal Frontiers in Plant Science

Received: *09 April 2022*
Accepted: *06 May 2022*
Published: *27 June 2022*

Citation:
Yang Z, Yang Y, Li C, Zhou Y, Zhang X, Yu Y and Liu D (2022) Tasseled Crop Rows Detection Based on Micro-Region of Interest and Logarithmic Transformation. Front. Plant Sci. 13:916474. doi: 10.3389/fpls.2022.916474

Machine vision-based navigation in the maize field is significant for intelligent agriculture. Therefore, precision detection of the tasseled crop rows for navigation of agricultural machinery with an accurate and fast method remains an open question. In this article, we propose a new crop rows detection method at the tasseling stage of maize fields for agrarian machinery navigation. The whole work is achieved mainly through image augment and feature point extraction by micro-region of interest (micro-ROI). In the proposed method, we first augment the distinction between the tassels and background by the logarithmic transformation in RGB color space, and then the image is transformed to hue-saturation-value (HSV) space to extract the tassels. Second, the ROI is approximately selected and updated using the bounding box until the multiple-region of interest (multi-ROI) is determined. We further propose a feature points extraction method based on micro-ROI and the feature points are used to calculate the crop rows detection lines. Finally, the bisector of the acute angle formed by the two detection lines is used as the field navigation line. The experimental results show that the algorithm proposed has good robustness and can accurately detect crop rows. Compared with other existing methods, our method's accuracy and real-time performance have improved by about 5 and 62.3%, respectively, which can meet the accuracy and real-time requirements of agricultural vehicles' navigation in maize fields.

Keywords: agricultural machinery navigation, crop rows detection, micro-region of interest, energy-efficient, logarithmic transformation

1. INTRODUCTION

In recent years, advances in intelligent agriculture have effectively reduced human costs and decreased the human harm caused by chemical factors such as pesticides. Real-time navigation of machines walking in the field is crucial for agriculture. Among them, the most popular approaches for field navigation are still path planning based on Global Position System (GPS) (Jin and Tang, 2011; Hameed, 2014; Li et al., 2019; Wang et al., 2021) and machine vision-based navigation (Ball et al., 2016; Radcliffe et al., 2018; Mavridou et al., 2019; Rovira-Mas et al., 2021; Vrochidou et al., 2022). The development of path planning algorithms has solved the path tracking problem of

agricultural machinery on a global scale, but the phenomenon of seedling injury from wheels is still inevitable. Since crops are usually sown in rows, machine vision-based field navigation is a promising way to provide navigation paths for agricultural machinery. Among them, the critical technology of computer vision, feature extraction (Manavalan, 2020; Xue et al., 2020, 2021a; Shrivastava and Pradhan, 2021; Vishnoi et al., 2022), is widely used in crop rows detection due to its advantages, such as low reliance on data resources. Many researchers have devoted significant efforts to developing efficient field navigation algorithms, and they can be classified into the following types.

1.1. Methods Based on Hough Transform

Hough (1962) proposed a way to transform points from a right-angle coordinate system into hough space. It has a good performance in processing information with straight-line features, but the real-time performance and accuracy of crop rows detection are unsatisfactory with mid-late stage plants. Thus, various improvements have been proposed. Ji and Qi (2011) detected crop rows by randomly selecting feature points for the Hough transform and then using many-to-one mapping to parameter space. Gall et al. (2011) established Hough Forest to improve the speed of Hough straight line detection. Winterhalter et al. (2018) proposed a reliable plant splitting pipeline and detected crop rows by Hough transform, but this approach is still limited to the crop rows at the early stage.

1.2. Methods Based on Horizontal Strips

It is very difficult to extract crop information from non-parallel crop rows in the image. This problem is effectively solved by dividing the image into multiple horizontal strips and processing them in successive steps. Ma et al. (2021) determined the number of crop rows by dividing horizontal strips in the upper part of the image. Ospina and Noguchi (2019) derived detection lines of crop rows by dividing horizontal strips. Crop contours in each strip are calculated, and their geometric centers are extracted as feature points for fitting. Zhou et al. (2021) determined the multi-ROI by dividing the horizontal strips. The initial ROI is calculated and continuously slides upward for the update. Finally, the midpoints of each ROI are fitted to make a navigation line. This method does not fully extract crop information when dealing with discrete characteristics of plants.

1.3. The Deep Learning Method

During the past decades, deep neural networks (DNNs) have made great success in field detection. Bah et al. (2020) combined Convolutional Neural Networks (CNN) and the Hough transform to detect crop rows in the field. Adhikari et al. (2020) used a deep network to learn semantic images, which makes the input images directly output detection lines as tractor control signals. Lac et al. (2022) first used a deep neural network to detect the plant stem and then used an aggregation algorithm to refine the localization of the crop further. Ponnambalam et al. (2020) divided the image into a vehicle driving area and a crop area using semantic segmentation based on CNN, and feature points are extracted. They further fitted the feature points derived from the multi-ROI to plan the crop rows detection lines. Although DNNs have good performance in accuracy, they really require large computing resources, and this limits their practical applications (Zhang et al., 2018a; Roy et al., 2019; Pan et al., 2021).

1.4. Integrated Approaches

Yu et al. (2021) proposed a treble classification and two-dimensional clustering-based crop rows detection in paddy fields for the problem of numerous weeds and floating weeds in the paddy fields. This method used a triple Otsu's (Otsu, 2007) method approach for segmentation and fitted the detection lines after selecting the misleading points by a two-dimensional adaptive clustering method. This method needs to be improved in terms of real-time performance. Jiang et al. (2015) integrated the crop rows with close geometric features in the robot walking area by multi-ROI for optimization and extraction of the crop rows centroids by clustering method. The detection lines were extracted by the linear regression method. Tenhunen et al. (2019) segmented the green objects after segmentation and obtained the direction and distance information between crop rows using a two-dimensional Fourier transform, then performed a clustering operation and finally obtained the location of the crop row. This algorithm is still deficient in coping with strong illumination conditions. Rabab et al. (2021) investigated adaptive crop row detection in variable field environments without the need to determine the number of crop rows by clustering. The method has good adaptability. Zhang et al. (2018b) defined clusters of feature points and fitted crop rows detection lines through a clustering algorithm and optimal path selection.

1.5. Our Contributions

In this study, we propose a new crop rows detection method for real-time navigation in maize fields during the tasseling stage. The article makes the following main contributions.

(1) To solve the difficulty of segmentation caused by the concentrated distribution of each color component in the image, we propose an image enhancement method based on logarithmic transformation, which well increases the contrast between the tassels and the background.

(2) We propose a method to determine ROI (Montalvo et al., 2012) by two steps of approximate selection and update, which overcomes the problem that ROI cannot achieve adaptivity in extracting information from skewed crop rows images.

(3) To verify the performance of the proposed method, various experiments are conducted to analyze the effect of parameters and make a comparison with the existing related studies. The experimental results demonstrate the advantages of this study in terms of accuracy and real-time performance.

Specifically, this study demonstrates the possibilities and prospects of feature extraction algorithms for extracting navigation lines in the tasseled maize field. We have overcome the problems of complex tassels segmentation and non-adaptive ROI. The whole process of this algorithm is shown in **Figure 1**. Our proposed model consists of two main parts: image preprocessing and determination of navigation lines. In image preprocessing: First, the logarithmic transformation is applied to augment this capture images. Second, these images are transformed into HSV space and segmented the tassels from

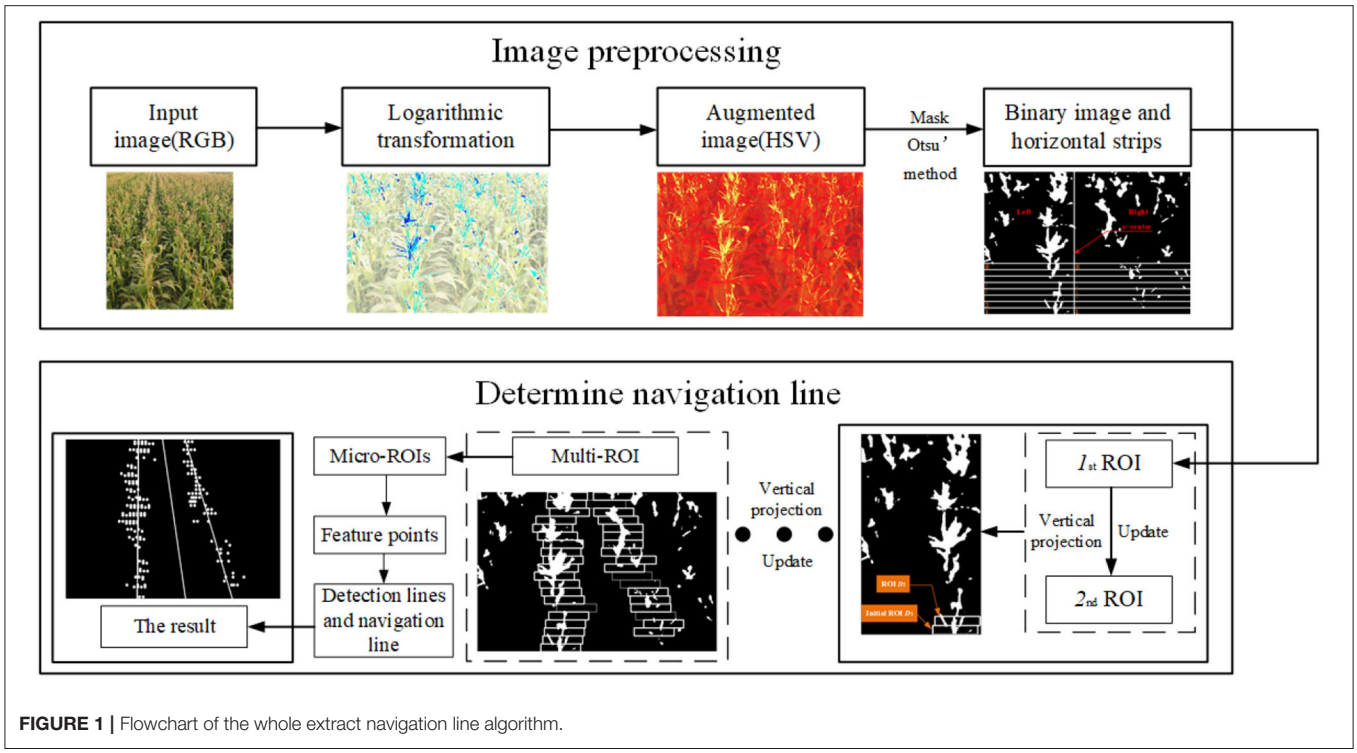


FIGURE 1 | Flowchart of the whole extract navigation line algorithm.

the background by grayscaling and Otsu's method. Finally, the binarised image is morphologically processed to decrease impulse noise. In the part of the determination navigation line: First, we divide the image into multiple horizontal strips, and then the initial ROI is determined using the bounding box. Second, continuously slide up the bounding box and update it until the whole image is completed to get the multi-ROI. Then, dividing the micro-ROIs to extract the feature points. Finally, the detection lines of the crop rows are made by using the least square method to fit these points. We further compute the bisector of the acute angle formed by the two detection lines and use it as the navigation line.

2. METHODS

While the deep learning-based methods achieve good performance in image processing and crop rows detection, they usually require a lot of computing resources, such as graphics processing units and GPU clusters. Based on image enhancement, selection of ROI, and delineation of horizontal strips, we propose an accurate and fast method for crop rows detection in the maize field during the tasseling stage. This section will introduce two main modules of our proposed method: image pre-processing and the determination of the navigation line.

2.1. Image Pre-processing

2.1.1. Pixel Value Modification

The distribution between crop rows in the upper part of the image is very concentrated, making it challenging to distinguish crop rows when extracting crop information. Additionally, the upper part of the image is not very meaningful for navigation; we only need the lower part of the image as the navigation area. Thus, the lower 3/4 of the image is taken as the processing object, and then the image is partially cropped to remove the redundant information. Finally, the size of the pixel value of the image is modified to 600×600 pixels.

2.1.2. Logarithmic Transformation

The grayscale processing (Liu et al., 2010, 2020; Laursen et al., 2014) is mainly performed using the excess green (ExG) (Comba et al., 2015; Tang et al., 2016; Chen et al., 2020) feature operator when segmenting the crops with the background at the early stage. However, ExG was not effective in processing maize at the tasseling location. The logarithmic function shows a nonlinear feature is uniformly increasing the independent variable. Additionally, the magnitude of change gradually decreases. Based on this, we adopt the logarithmic transformation for the pixel values of each pixel point in the image under RGB color space, which can augment the distinction between the tassels and background and realize the segmentation of the picture. When establishing the logarithmic function, the following factors are considered: (1) Transformation should avoid negative results after the logarithmic operation; (2) The effect of differentiation between the tassels and the background is augmented after the logarithmic operation. The logarithmic function established is Equation (1).

$$Out = C \times log(1.0 + p), \tag{1}$$

where *Out* is the result of the pixel value operation of the pixel point; *p* is the pixel value of the pixel point in the image to be processed; and *C* is a constant.

2.1.3. Building Masks and Morphological Processing

The images calculated in the way of 2.1.2 are transformed from RGB to HSV color space. We will determine the suitable threshold to build the mask through subsequent experiments. The mask images are converted to grayscale images, and Otsu's method (Cellini et al., 2017) is performed to extract the tassels. At the end of this part, it is necessary to select the appropriate kernels for the morphological processing of the image.

2.2. Determination Navigation Line

In this section, we describe how to determine a navigation line. It mainly contains two parts: selecting ROIs and planning navigation lines.

2.2.1. Select ROIs

The algorithm addresses the problem of navigation line extraction for field vehicles. We believe that the crop rows in the traveling area (the two crop rows in the center of the image) are valid for navigation, while the crop rows at the image boundary can be disregarded. Due to the perspective principle, the crop rows are not parallel in the image, bringing more significant difficulties to feature point extraction. Therefore, we extracted feature points by selecting ROIs.

To determine the ROI of an image, we specify a bounding box, which is described as follows: The coordinates of point q_0 *(Xr, Yr)* are used as the origin, *L1*+*L2* (*L1*=45pix, *L2*=55pix) as the width and *H* (20pix) as the height to determine the bounding box shown in **Figure 2**. Since *L1*, *L2*, and *H* are all constants and only the coordinates of the center point q_0 are variable. Thus, the location of the bounding box is expressed in Equation (2).

$$B = (X_r, Y_r). \tag{2}$$

2.2.1.1. Divide the Image Band

We divide the binarized images according to the following way: The image is divided into left and right areas using the line *x*=center (260pix) as the dividing line. The areas are labeled as *Left* and *Right*. The algorithm uses the same approach for *Left* and *Right*, thus, we only describe the process of *Left* in the following step. On *Left*, eight horizontal strips are divided in step length of Δh (20pix) from bottom to top, each strip was numbered *Ks* (s=1,2,3...,8). The resulting model is shown in **Figure 3**.

The unstable environment such as light and weeds makes the work hard of getting crop rows information, we determine the ROI by approximate selection and update. The primary choice is completed first. The bounding box is applied to frame the approximate position of the initial ROI. Then this ROI is updated according to its local pixel distribution to obtain a more accurate result of the initial ROI.

We start by approximately selecting the initial ROI with a bounding box, marking it as B_1, and the process is as follows: First, we set the threshold *Y=15*. The cumulative value *M(j)* of the number of white pixels in each column of the strip is calculated

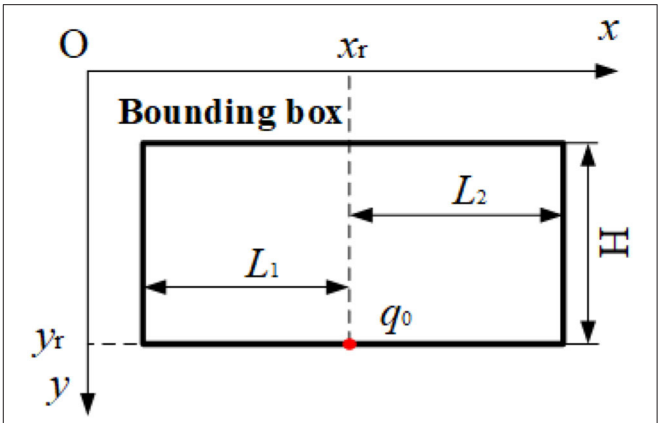


FIGURE 2 | The structure of the bounding box.

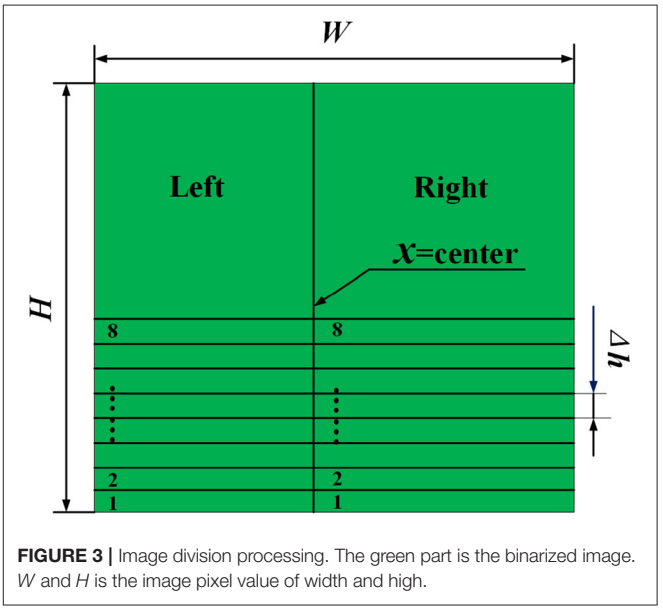


FIGURE 3 | Image division processing. The green part is the binarized image. *W* and *H* is the image pixel value of width and high.

using Equation (3) sequentially in the order of the labels until the maximum value of $M(j) \geq Y$ in the *Ks* strip, and this strip is name K_j. The values of columns whose coordinates are higher than *Y* are counted, and it is stored in list *Q*. In *Q*, the closest value to the *center* is the horizontal coordinate of the origin of the initial bounding box, and the lower boundary of the strip is the vertical coordinate of the origin of the initial bounding box. The coordinates of the origin of the bounding box are given by Equation (4), then B_1 is framed.

$$M(j) = \frac{\sum_{i=1}^{\Delta h} p_{(i,j)} = 255}{255}, j \in \{1, 2, ..., w\}, \tag{3}$$

where *j* is the column coordinate of the pixel point of the image strip; *i* is the row coordinate of the pixel point of the image strip; *p(i,j)* is the pixel value of the coordinate; *w* is the horizontal pixel

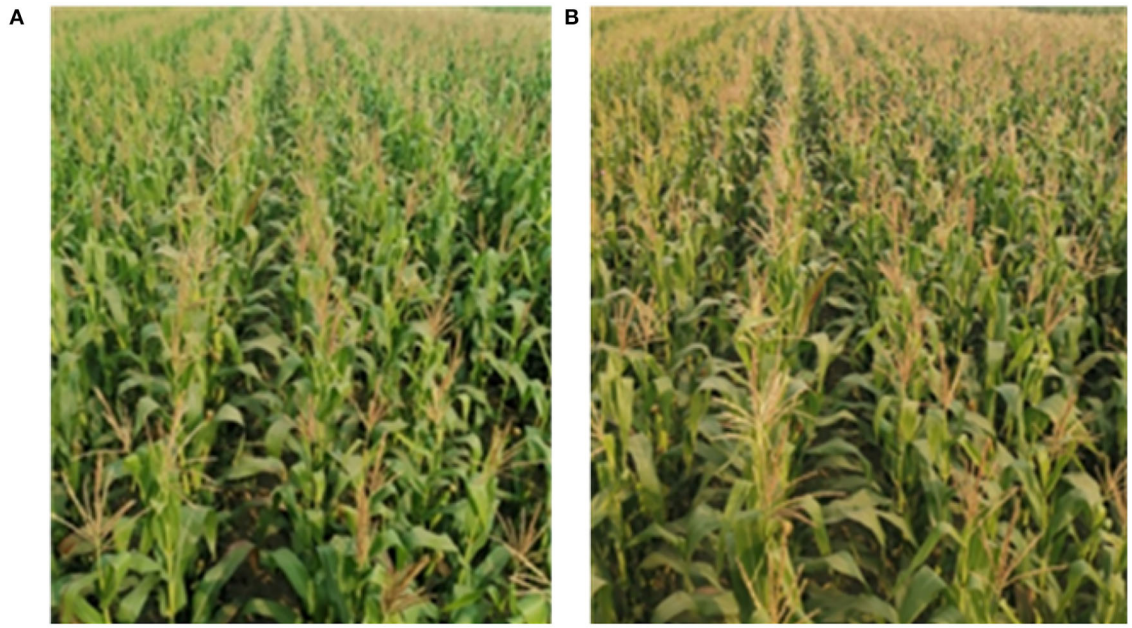


FIGURE 4 | Image captured by CMOS. **(A)** Type-1 field (illumination: 130800 lx). **(B)** Type-2 field (illumination: 101700 lx).

size of the strip.

$$(j_s, W - (K_s - 1) \times \triangle h). \tag{4}$$

2.2.1.2. The Method of Updating Bounding Box

To achieve a more accurate location of the ROI, we propose an updated method, and the steps are as follows.

(1) First, the number of white pixels in each column of the bounding box is counted using Equation (5). Mark the horizontal coordinate of the lower boundary of the bounding box as x_p (x_p=1,2,3..., *L1+L2*). The distribution of white pixels in the bounding box for the part of the horizontal coordinate less than x_p and more than x_p are expressed respectively as Equation (6) and Equation (7).

$$Z(v) = \frac{\sum_{u=1}^{\Delta h} p_{(u,v)} = 255}{255}, v \in \{1, 2, ..., L1 + L2\}, \tag{5}$$

where u is the row coordinate of the pixel point within the ROI; v is the column coordinate of the pixel point within the ROI; $p_{(u,v)}$ is the pixel value of this point.

$$I_l = \sum_{v=1}^{x_p} Z(v) \times (x_p - v\,), \tag{6}$$

$$I_r = \sum_{v=x_p+1}^{L1+L2} Z(v) \times (v - x_p\,), \tag{7}$$

where *Z(v)* is calculated by Equation (5) v is the column coordinate of the pixel point of the image strip; *L1+L2* is the width pixel value of the bounding box.

(2) To find a point in the bounding box that makes the distribution of pixels balanced, we build an evaluation function $f(x_p)$, which is shown as Equation (8). For the second-order partial derivative (Equation (9) of $f(x_p)$. Since Equation (9) >0, there is a minimum value of $f(x_p)$. The horizontal coordinates of the bounding box are updated to the horizontal coordinates of the lowest point of $f(x_p)$ which is determined by Equation (10), and the vertical coordinates are unchanged.

$$f(x_p) = |I_l - I_r|\,, \tag{8}$$

$$\frac{\partial^2 f(x_p)}{\partial x_p^2} = 2, \tag{9}$$

$$j_k = argminf(x_p). \tag{10}$$

2.2.1.3. Determining the Final Result of the Initial ROI

By means of 2.2.1.2 [Due to the white pixels of B_1 are already calculated by Equation (3) and do not need to be repeated by Equation (5)], the bounding of B_1 is updated, and the coordinate origin of the result is $(j_{k1}, W\text{-}(K_s\text{-}1)\times\Delta h)$ [j_{k1} is the result of the calculation of Equation (10)]. This bounding box determines the final outcome of the initial ROI, marked as D_1.

2.2.1.4. Sliding of Bounding Box

After obtaining the initial ROI, the ROI named B_2 is determined by sliding the bounding box of the initial ROI in the negative direction along the y-axis in step length Δh. The coordinate of the

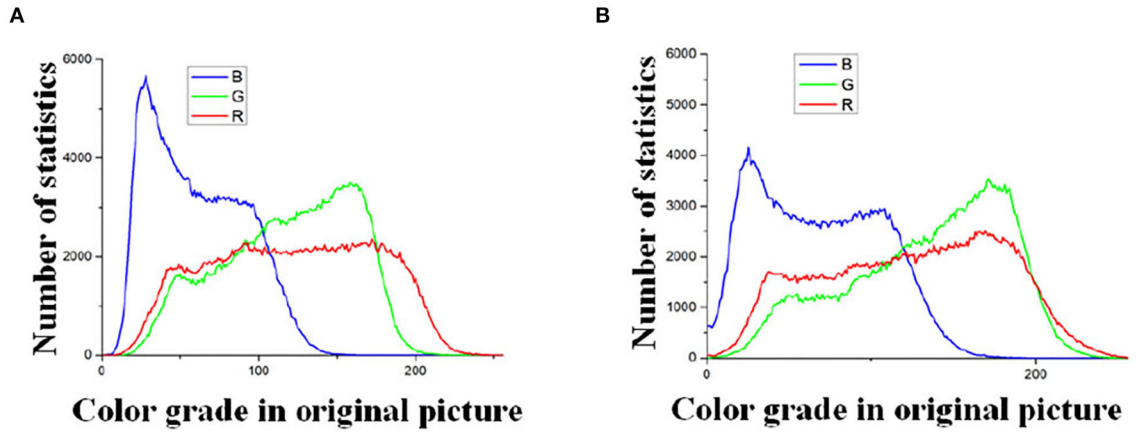


FIGURE 5 | Statistics of color components in RGB space. **(A)** Color grade in original picture Type-1 field (illumination: 130800 lx). **(B)** Color grade in original picture Type-2 field (illumination: 101700 lx).

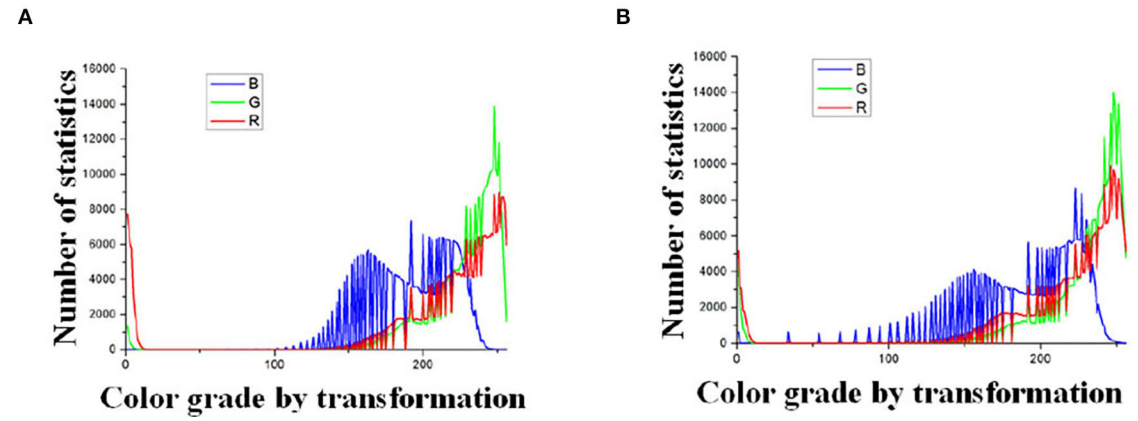


FIGURE 6 | Statistics of color components after logarithmic transformation. **(A)** Color grade by transformation Type-1 field (illumination: 130800 lx). **(B)** Color grade by transformation Type-2 field (illumination: 101700 lx).

bounding box origin of B_2 is $(j_{k1}, W\text{-}K_s \times \Delta h)$. T_1 is the number of white pixel points inside B_2, calculated by Equation (11).

$$T_1 = \sum_{n=1}^{L1+L2} \sum_{b=1}^{\Delta h} p(b,n) = 255, \tag{11}$$

where b is the row coordinate of the pixel point; n is the column coordinate of the pixel point. $p(b,n)$ is the value of the pixel point in B_2.

The algorithm updates the bounding box of the optimized B_2 to obtain the updated result D_2 by establishing a threshold T_0. We take T_0 as 20 and offset d=20 in this article. The updated result D_2 for B_2 is derived and framed. The updating method is as follows:

(1) $T_1 < T_0$: B_2 is judged to be a sparse crop area. Considering that the distribution of the crop rows in the image is skewed, we slide the bounding box of B_2 in the image coordinate system along the x-axis toward the line x=center by sliding step of offset d, and the result of D2 is determined.

(2) $T_1 \geq T_0$: B_2 is determined as a feature region. The bounding box of B_2 is updated by the method of 2.2.1.3 to give the coordinates of the origin as $(j_{k2}, W\text{-}K_s \times \Delta h)$, and the result frame D_2 [j_{k2} is the result of the calculation of Equation (10)].

2.2.1.5. Determination of Multi-ROI

There is a tendency for the crop rows to converge in the image. Based on this feature, the steps in 2.2.1.4 are repeated. The process is as follows. First, The bounding box of the current ROI(D_e) (subscript e is the number of the ROI, e=1,2,3...) is slid up to determine the subsequent ROI (B_{e+1}) in the strip. Then B_{e+1} is updated and optimized to obtain. D_{e+1} until the multi-ROI is reached.

2.2.2. Planning Navigation Line by Getting Micro-ROIs

The part marked as *Right* are processed using the methods in 2.2.1.1, 2.2.1.3, 2.2.1.4, and 2.2.1.5 in sequence. Therefore, all the ROI of the whole picture are determined.

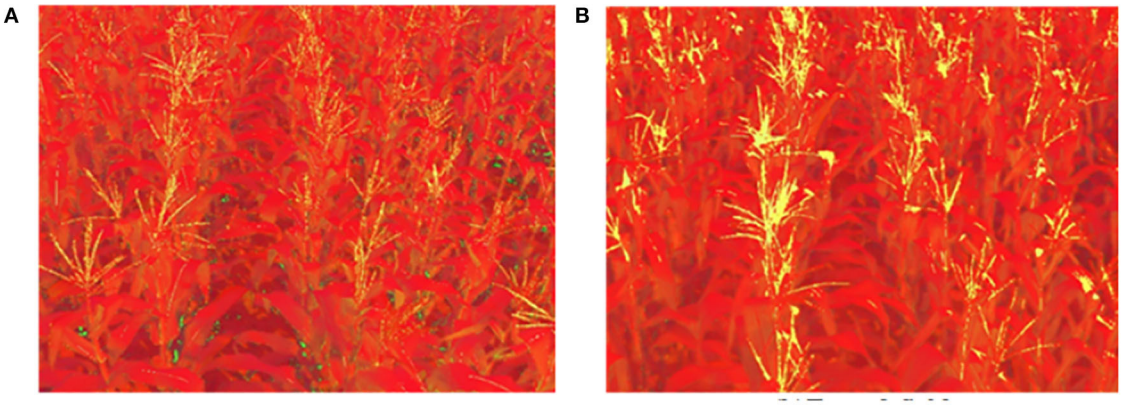


FIGURE 7 | Images in HSV color space. **(A)** Type-1 field (illumination: 130800 lx). **(B)** Type-2 field (illumination: 101700 lx).

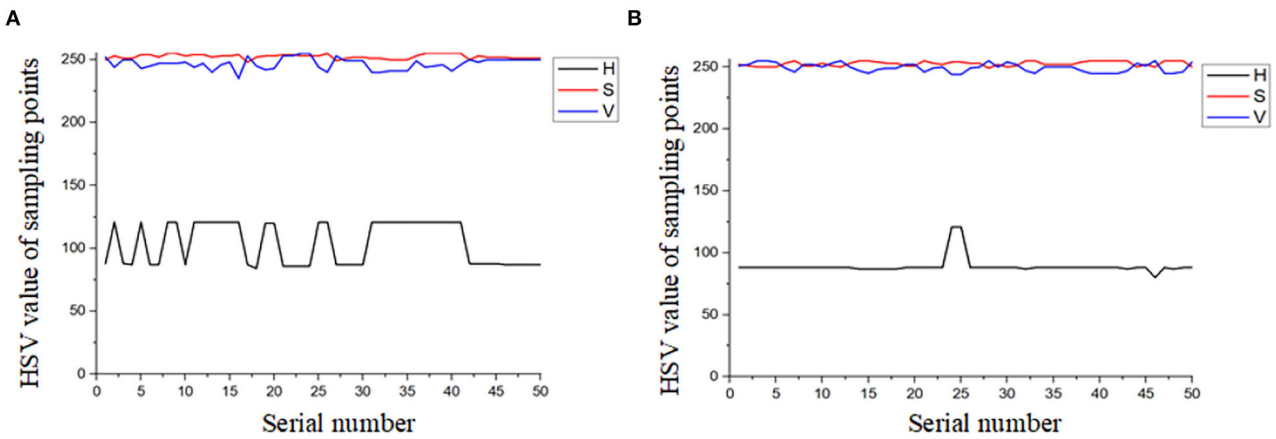


FIGURE 8 | Distribution interval of H, S, and V at sampling points. **(A)** Type-1 field (illumination: 130800 lx). **(B)** Type-2 field (illumination: 101700 lx).

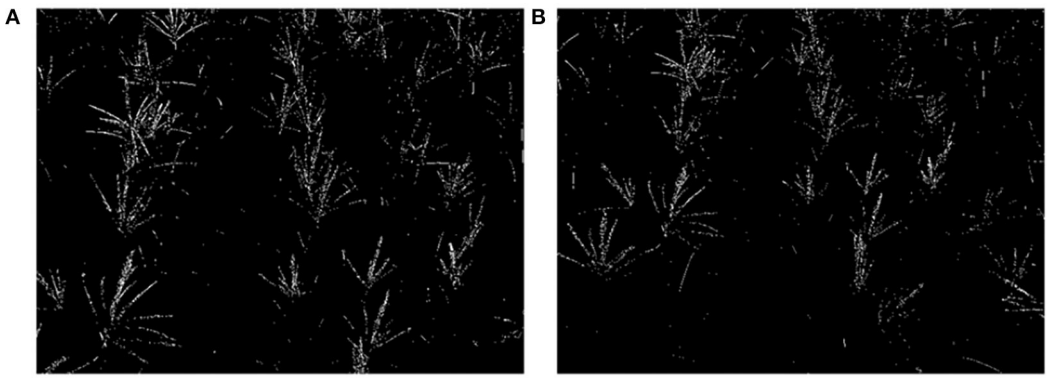


FIGURE 9 | Binary images. **(A)** Type-1 field (illumination: 130800 lx). **(B)** Type-2 field (illumination: 101700 lx).

This algorithm extracts feature points by building micro-ROIs. The method is as follows: Each ROI (D_1, D_2,...) is divided into 10 x 2 grids, each with a 10(pix)×10(pix) micro-ROI. We set the threshold for the number of white pixels in the micro-ROI H_0=20 and calculate the number of white pixels H_n in the individual micro-ROI. When

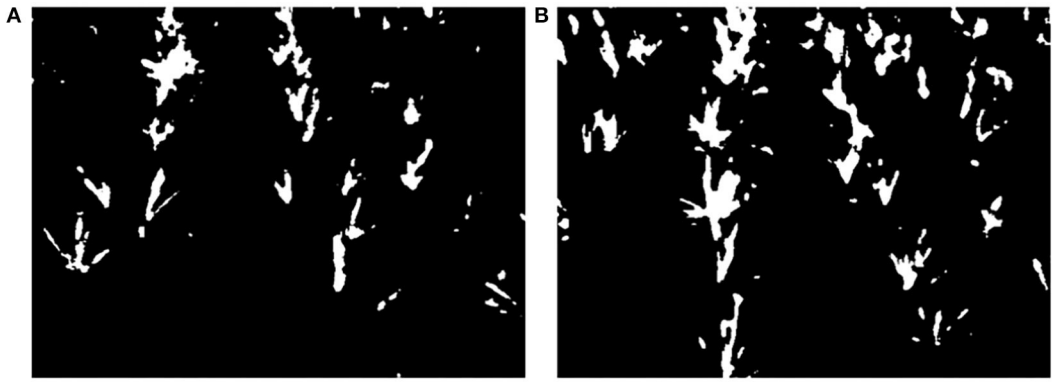


FIGURE 10 | Results of morphological processing. **(A)** Type-1 field (illumination: 130800 lx). **(B)** Type-2 field (illumination: 101700 lx).

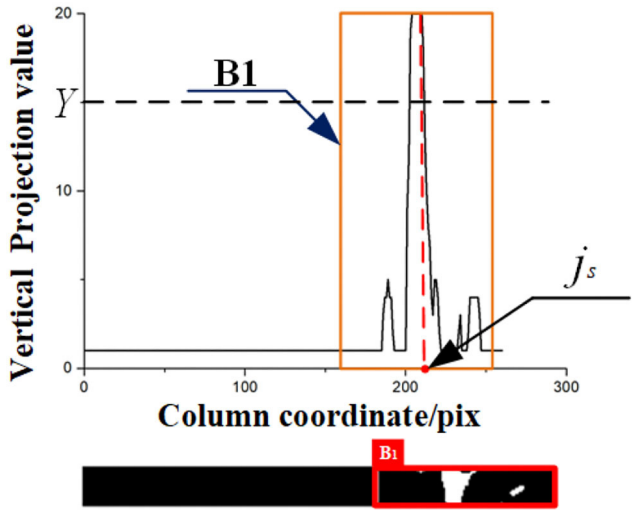


FIGURE 11 | The projected image of the K_j strip (K_j is calculated in 2.2.1.1).

$H_n > H_0$, the geometric midpoint of the micro-ROI is a feature point.

Finally, these feature points are fitted using least squares to get the crop rows detection lines. The angle bisector at the intersection of the two identification lines is used as the navigation line.

3. EXPERIMENTAL RESULT AND DISCUSSION

In this section, we first introduce image acquisition. Then the results of each step are described. Finally, the performance of this article and existing algorithms is shown.

3.1. Image Acquisition and Processing Equipment

The subject of this study is the images of maize crops in the tasseling stage. The image acquisition device is a CMOS (Complementary Metal-Oxide-Semiconductor) machine vision camera, which is installed at a height of 2.9 m from the ground with a tilt angle of 20° the ground and calibrated using the camera imaging principle. The camera resolution is 1,920 × 1,080 pixels. Video is collected in Gu'an County, Hebei Province, China. The crops are planted at a row spacing of 60 cm, and the plant height was 2.7 m. The programming software was python 3.7 IDE, PyCharm professional 2020 compiler. The image processing hardware used Advantech MIC-7700 IPC, processor Intel Corei5, main frequency 2.5GHz, graphics card for NVIDIA GTX 1650, video memory 8G. The video was saved in AVI format. The video was collected on 8 July 2020 (illumination:101700 lx) and 11 July 2020 (illumination:130800 lx).

3.2. The Performance of the Proposed Model

In this part, we will use these images as examples to describe the result of every step. Example images are shown in **Figure 4**. In **Figure 4**, the illumination of the Type-1 field and Type-2 field are 130800 (lx) and 101700 (lx) respectively. First, the pixels of the two images are modified by way of 2.1.1. Then the results are processed in the following steps.

3.2.1. The Result of Logarithmic Transformation

The processing results of **Figure 4** are plotted with the frequency statistics of each color component (R, G, B) as shown in **Figure 5**. It can be found that the distribution of each color component is concentrated, indicating that the differentiation between the tassels and the backgrounds, such as leaves and soil, is not obvious under natural illumination, which makes it challenging for subsequent segmentation work. To determine the value of C in Equation (1), the experiment methods are as follows: We transformed the images by Equation (1), and the results of C taking values in the range 1 to 100 were observed. We found that the frequency distribution has clear discrimination when C is taken as 48. The statistical results are shown in **Figure 6**. Comparing the statistical results in **Figures 5**, **6**, the application of logarithmic transformation significantly enhances the distinction between the tassels and the background, which

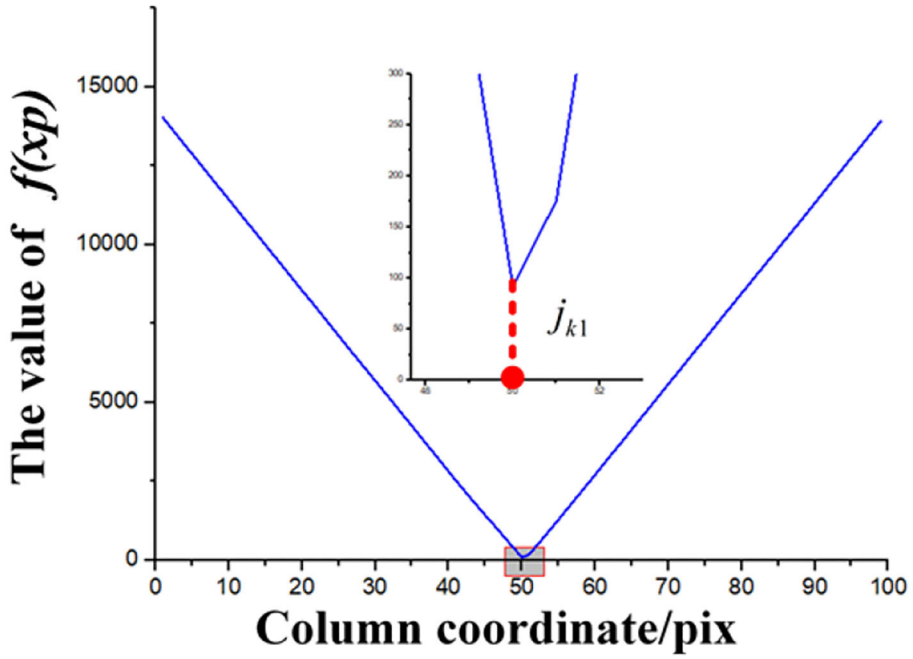


FIGURE 12 | Distribution of the judging function.

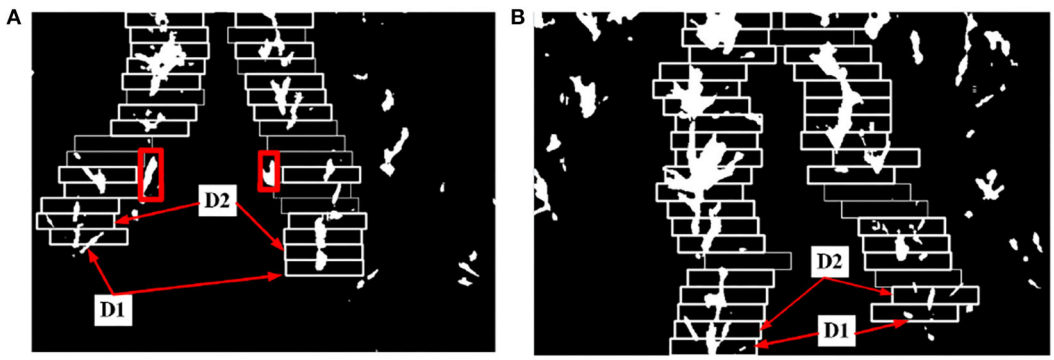


FIGURE 13 | The result of multi-ROI extraction. **(A)** Type-1 field (illumination: 130800 lx). **(B)** Type-2 field (illumination: 101700 lx).

provides the necessary conditions for the subsequent image segmentation, so we determine the value of C is 48.

3.2.2. Result of Masks and Morphological Operations

Images are transformed to HSV color space, and the result is shown in **Figure 7**. Fifty frames are randomly selected, and 50 sampling points are chosen on each of the tassels to count the distribution of H, S, and V. The statistical results are shown in **Figure 8**. Based on the statistical results, we set the threshold of H as [80,121], S as [250, 255], and V as [240, 255]. By these thresholds, the masks are built, mask images are transformed into grayscale images, and Otsu's method is performed to extract the tassels. The result is shown in **Figure 9**.

There is impulse noise in the binary images, and the feature pixel values of maize tassels are too small, making it challenging to extract feature points. It is necessary to perform morphological operations. Thus, we use a convolutional kernel of size 3×3 to inflate the image once morphologically. Then use a median filter with a convolutional kernel of size 9 to the noise for the inflated image. The morphological processing results are shown in **Figure 10**.

3.2.3. The Result of Dividing Image, Initial ROI, and Second ROI

By the method of 2.2.1.1 to 2.2.1.3, the result of every step is as follows: Take **Figure 10B** as an example. We find the k_s=1 horizontal strip. Its statistical graph of the number of white pixel dots in the column is shown in **Figure 11**. With the assistance of this graph, the rough initial ROI $B1$ is determined, and the

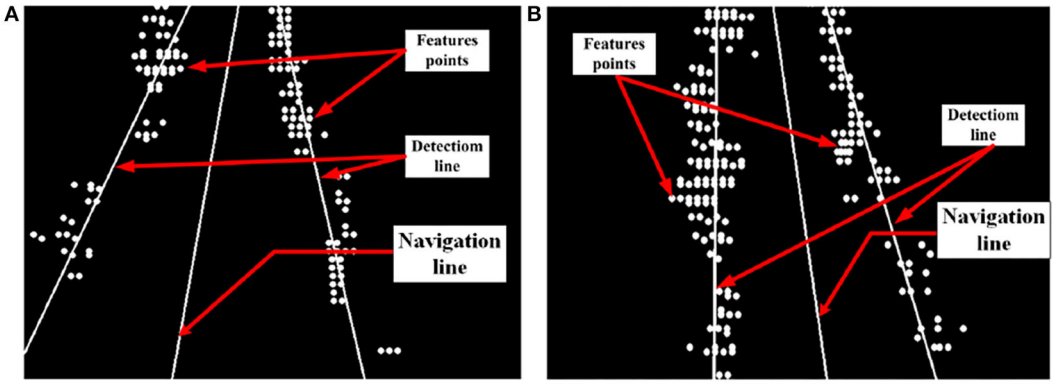


FIGURE 14 | Fitting results for feature points. **(A)** Type-1 field (illumination: 130800 lx). **(B)** Type-2 field (illumination: 101700 lx).

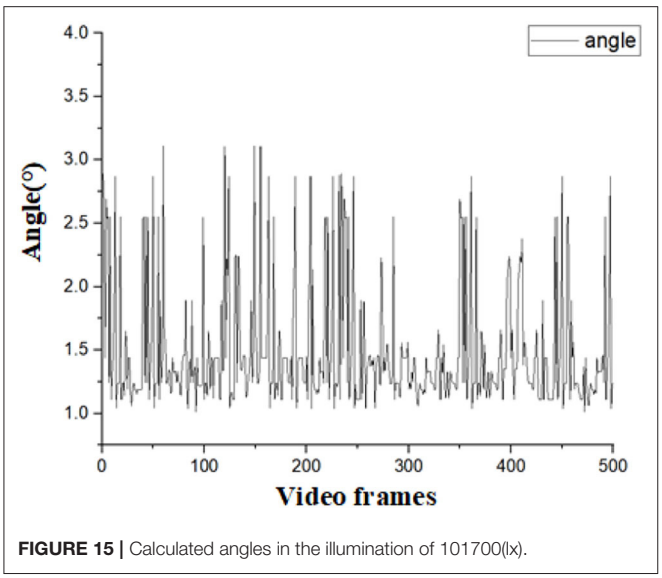


FIGURE 15 | Calculated angles in the illumination of 101700(lx).

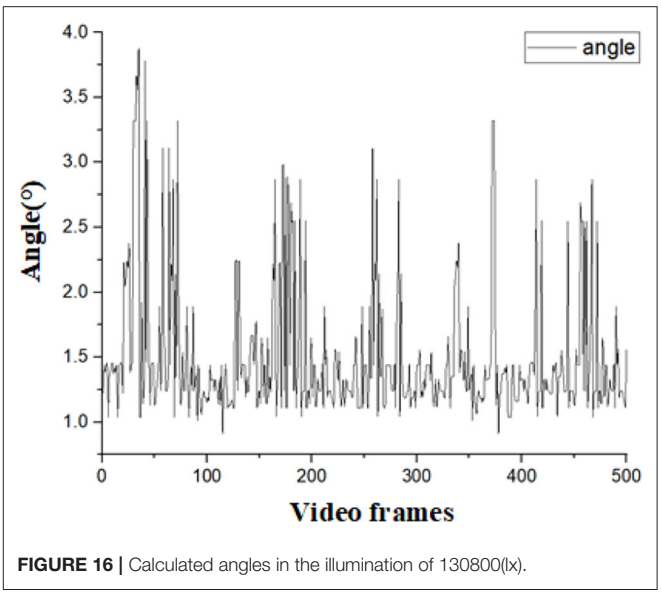


FIGURE 16 | Calculated angles in the illumination of 130800(lx).

result is shown in **Figure 11**. By intercepting the curve inside B_1 in **Figure 11**, we plot the distribution function of $f(x_p)$ as shown in **Figure 12**, and find the horizontal coordinate of the lowest point is j_{k1}. The coordinates of the origin point of the bounding box are $(j_{k1}, W\text{-}(K_s\text{-}1)\times\Delta h)$. We use the bounding box to frame the final result of the initial ROI D_1. Next, the second ROI can be determined by the method proposed in 2.2.1.4. The results are shown in **Figure 13**.

3.2.4. The Result of Multi-ROI

The final multi-ROI can be determined by the method proposed in 2.2.1.5, the result is shown in **Figure 13**. The tassel feature pixels that deviate from the path appear during the upward sliding of the bounding box, as shown in the red box in **Figure 13A**. According to the original images, the pixels in this region correspond to tassels deviated from the crop rows, so this region is judged as deviated and will not be processed. Our proposed method allows the ROI always to follow the crop rows trend and guarantees the reliability of the subsequent feature point extraction work as much as possible.

TABLE 1 | The performance of different methods in indicators of time,error angle, accuracy, and FPS compared with those in the literature.

Methods	Average time(s)	Maximum angle(θ)	Minimum angle(θ)	Average angle(θ)	Accuracy (%)	FPS
This study	312.3	3.88	1.04	1.49	98.6	4.4
Algorithm-1	1000.2	12.82	8.69	8.36	70.1	2.1
Algorithm-2	828.4	9.24	5.88	6.89	88.5	3.8
Algorithm-3	459.9	4.95	3.28	3.78	93.6	2.9
Algorithm-4	610.4	10.4	7.77	8.59	84.9	2.6

3.2.5. The Result of Feature and Navigation Line

By the method proposed in 2.2.2, feature points are extracted. We further plan the crop rows detection lines and the field navigation line. The result is shown in **Figure 14**.

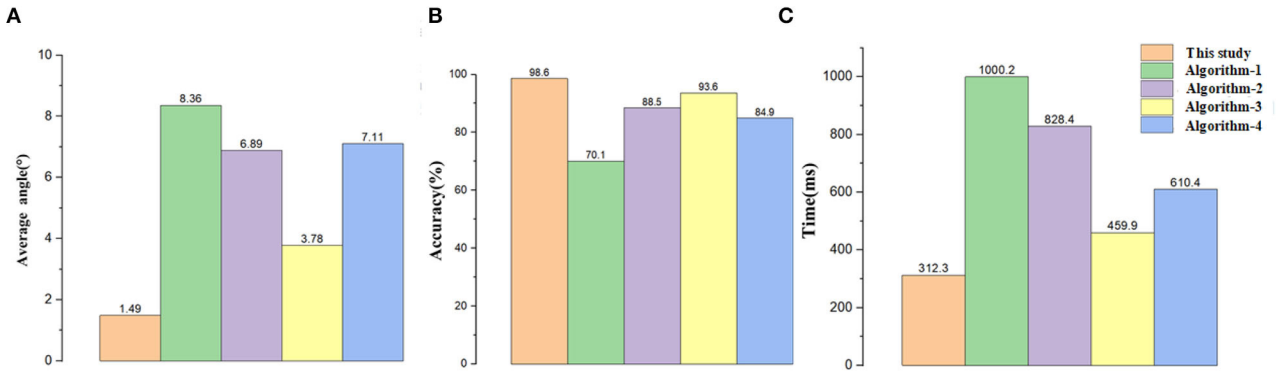


FIGURE 17 | Results of comparative data. **(A)** The data of angle, **(B)** the data of accuracy, and **(C)** the data of time.

3.3. Results and Discussion

To verify the accuracy of the detection results. We designed an experiment as follows: Since the navigation line is calculated by crop rows detection lines, only the declination angle θ between the navigation line and the manually drawn navigation lines is needed to evaluate the detection accuracy of the algorithm. If θ is less than 5°, the detection result for navigation can be considered accurate. As the speed of agricultural machinery in the field is very slow, no more than 0.5m/s, we considered that the algorithm can meet the basic real-time requirements when the frames per second (FPS) is greater than 4. A total of 1,000 frames of video taken in the field were randomly selected for the experiment. Among this video, 500 frames are in 101700(lx) illumination, and others are in 130800(lx) illumination. Average processing time, θ, and FPS were calculated for each frame selected. The error angles in different illumination are shown in **Figures 15**, **16**, respectively. The performance of this algorithm is shown in **Table 1**.

To further verify the reliability and real-time performance of this study it is compared with the methods proposed by Hough (1962) (Hough Transform) (Algorithm-1), Ji and Qi (2011) (Algorithm-2), Zhou et al. (2021) (Algorithm-3), and Zhang et al. (2018b) (Algorithm-4). Additionally, we will analyze the results of the performance among different methods in terms of accuracy and processing time. The Hough transform is a very classical algorithm. After the images are segmented using the pre-processing method proposed, we used the Hough transform to detect the crop rows and record the data. Algorithm-2 mapped the coordinates of the image space to the parameter space through random numbers. The parameter space feature points disappeared when dealing with the case of divergence of tassels, resulting in significant deviations in the Hough line detection process, and the navigation line's average error angle reached 6.89°. Moreover, Algorithm-2 used the Hough transform algorithm, which had a higher computational cost, and the average time to process a frame was more than 800 ms, which could not meet the real-time requirements of field navigation. In response to the problems of Algorithm-2, the proposed algorithm took the left and suitable regions of the image and obtained the multi-ROI by sliding off the bounding box to extract the feature points, which effectively calculated the navigation line under the dense conditions of crop rows distribution. The average error angle of the navigation line calculated by Algorithm-3 is 3.78°. This error is because the ROI window proposed in the paper could not completely cover the tassels after the ROI window was slid upward when the tassels were bifurcated during the extraction process, which led to deviations in the fitting. Furthermore, the algorithm selects the appropriate feature points by the center of each ROI. However, the navigation lines were biased due to the problems of forked tassels. The algorithm in this article deals with the deviation points by fixed-size ROI and the ROI optimization to extract the tassels completely. Algorithm-4 removed feature points by selecting the midpoints of the left and right edge points in the image strips. This approach had better performance when dealing with small target plants, but when dealing with more oversized tassels, especially the characteristic tassels with bifurcation, the feature points calculated in this way did not express the crop distribution in this region, resulting in a significant error in the subsequent detection lines extraction. The proposed algorithm extracted feature points more completely by dividing micro-ROIs, avoiding the above problems. By comparison, the algorithm of this study has high accuracy and real-time performance in crop row detection in maize fields at the tasseling stage.

The performance of each algorithm is shown in **Table 1**. The main parameters are shown in **Figure 17** so that they can be more intuitive. Compared to current popular algorithms, we have improved accuracy by at least 5% and single-frame processing time by at the least 62.3%. The processing results of this algorithm are shown in **Figure 18**. The results of the above comparison experiment are shown in **Figure 19**.

4. CONCLUSION AND FUTURE STUDY

Based on machine vision, we propose a real-time method for the extraction of navigation lines in the maize field during the tasseling stage. Field navigation line extraction in the maize crop rows during the tasseling phase is challenging because the tassels are hard to be segmented from the background, and extracting their information is difficult. We propose a real-time crop rows

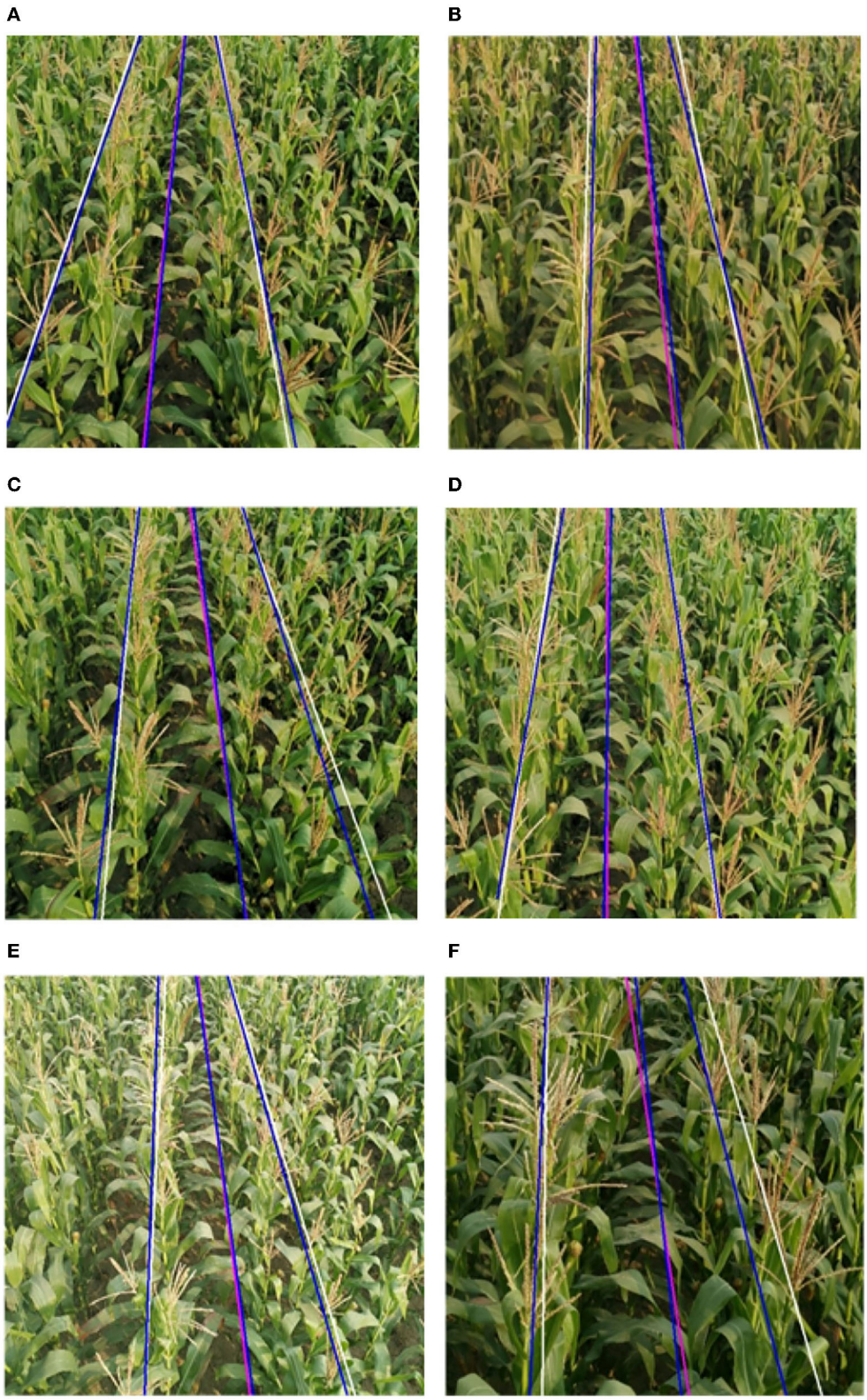


FIGURE 18 | Experimental results of crop rows detection. The blue lines are the manual calibration lines, the white lines are the algorithmic crop rows identification lines, and the pink lines are the algorithmic navigation lines. **(A)** Type-1 field, **(B)** Type-2 field, **(C)** Type-3 field, **(D)** Type-4 field, **(E)** Type-5 field, and **(F)** Type-6 field.

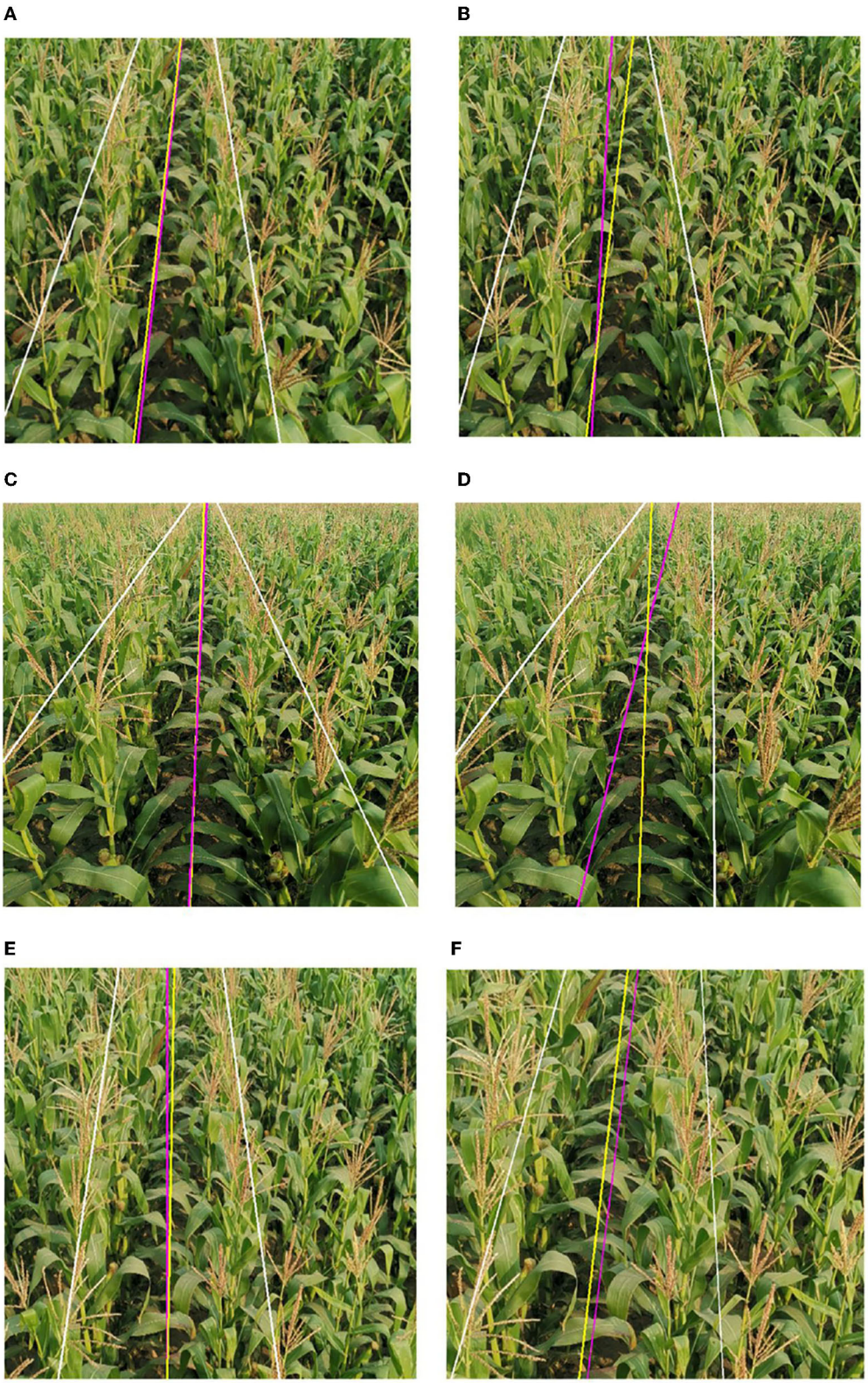


FIGURE 19 | Extraction results comparison between literature and algorithm in this article. This algorithm's crop rows detection lines are white, the manually calibrated navigation line is yellow, and the corresponding algorithm's navigation line is pink. **(A)** Algorithm of this paper, **(B)** Algorithm-2, **(C)** Algorithm of this paper, **(D)** Algorithm-3, **(E)** Algorithm of this paper, and **(F)** Algorithm-4.

detection algorithm based on logarithmic transformation and micro-ROI in the field during the tassel period to address these issues. After cutting the image captured by the CMOS camera (RGB) to obtain a 600×600 (pix) image, the algorithm performs a logarithmic transformation to augment the distinction between the tassel and the background. This research converted the image to the HSV color space. Additionally, the background of the tassels was created as a mask. After the lower part of the image is divided into eight horizontal strips, we use the bounding box to determine the initial ROI by selecting the starting point. The final result of the initial ROI is determined by updating its bounding box position. The current bounding box is slid along the negative direction of the y-axis of the image coordinate system in steps Δh and updated until the multi-ROI is reached. The ROI is divided into cells to get the micro-ROIs, and then the feature points are extracted. Feature points are fitted to derive crop row detection lines, on which the navigation lines are then calculated. We fit the feature points to make crop rows detection and navigation line. The error in the navigation line of this algorithm is stable at 1.49°, and the average computational time of the single frame is 312.3 ms. The accuracy is reaching 98.6%. After the Comparison experiment, the algorithm proves to have a clear advantage in terms of real-time and accuracy.

However, there are still some limitations to this method, such as different climates and different crop row spacing, which can reduce the accuracy of the algorithm. In the future, we will focus on new methods of feature extraction (Xue et al., 2021b), image augments (Sui et al., 2020), and ROI adaptability to segment a variety of tasseled plants and calculate adaptive ROI with a wide range of planting rows. In addition, how to apply the energy-efficient spiking neural networks to crop rows detection is another interesting topic. Because the SNNs hold the potential to provide a good performance equivalent to that of DNNs while with low latency and high computational efficiency (Feldmann et al., 2019; Zhang et al., 2021; Luo et al., 2022).

DATA AVAILABILITY STATEMENT

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

AUTHOR CONTRIBUTIONS

YZ and YYa performed all the experiments. All authors contributed to the design of the experiment, result interpretation, and writing.

FUNDING

The authors acknowledge that the research was financially supported by the Nation Key Research and Development Program of China (Grant No. 51905004), the University Synergy Innovation Program of Anhui Province (Grant No. GXXT-2020-011), Central Leading Local Science and Technology Development Special Foundation for Sichuan Province PR China (No. 2021ZYD0020), and Scientific Research Project of Yibin University (No. 0219024502).

REFERENCES

Adhikari, S. P., Kim, G., and Kim, H. (2020). Deep neural network-based system for autonomous navigation in paddy field. *IEEE Access* 8, 71272–71278. doi: 10.1109/ACCESS.2020.2987642

Bah, M. D., Hafiane, A., and Canals, R. (2020). Crownet: deep network for crop row detection in uav images. *IEEE Access* 8, 5189–5200. doi: 10.1109/ACCESS.2019.2960873

Ball, D., Upcroft, B., Wyeth, G., Corke, P., English, A., Ross, P., et al. (2016). Vision-based obstacle detection and navigation for an agricultural robot. *J. Field Rob.* 33, 1107–1130. doi: 10.1002/rob.21644

Cellini, A., Blasioli, S., Biondi, E., Bertaccini, A., Braschi, I., and Spinelli, F. (2017). Potential applications and limitations of electronic nose devices for plant disease diagnosis. *Sensors* 17, 2596. doi: 10.3390/s17112596

Chen, J. Q., Qiang, H., Xu, G. W., Liu, X., Mo, R. X., and Huang, R. Z. (2020). Extraction of navigation line based on improved grayscale factor in corn field. *Ciencia Rural* 50, 699. doi: 10.1590/0103-8478cr20190699

Comba, L., Gay, P., Primicerio, J., and Aimonino, D. R. (2015). Vineyard detection from unmanned aerial systems images. *Comput. Electron. Agric.* 114, 78–87. doi: 10.1016/j.compag.2015.03.011

Feldmann, J., Youngblood, N., Wright, C. D., Bhaskaran, H., and Pernice, W. H. (2019). All-optical spiking neurosynaptic networks with self-learning capabilities. *Nature* 569, 208–214. doi: 10.1038/s41586-019-1157-8

Gall, J., Yao, A., Razavi, N., Van Gool, L., and Lempitsky, V. (2011). Hough forests for object detection, tracking, and action recognition. *IEEE Trans. Pattern Anal. Mach. Intell.* 33, 2188–2202. doi: 10.1109/TPAMI.2011.70

Hameed, I. A. (2014). Intelligent coverage path planning for agricultural robots and autonomous machines on three-dimensional terrain. *J. Intell. Rob. Syst.* 74, 965–983. doi: 10.1007/s10846-013-9834-6

Hough, P. (1962). *Method and Means for Recognizing complex Patterns*. U.S.Patent.

Ji, R. H., and Qi, L. J. (2011). Crop-row detection algorithm based on random hough transformation. *Math. Comput. Model.* 54, 1016–1020. doi: 10.1016/j.mcm.2010.11.030

Jiang, G. Q., Wang, Z. H., and Liu, H. M. (2015). Automatic detection of crop rows based on multi-rois. *Expert Syst. Appl.* 42, 2429–2441. doi: 10.1016/j.eswa.2014.10.033

Jin, J., and Tang, L. (2011). Coverage path planning on three-dimensional terrain for arable farming. *J. Field Rob.* 28, 424–440. doi: 10.1002/rob.20388

Lac, L., Da Costa, J. P., Donias, M., Keresztes, B., and Bardet, A. (2022). Crop stem detection and tracking for precision hoeing using deep learning. *Comput. Electron. Agric.* 192, 106606. doi: 10.1016/j.compag.2021.106606

Laursen, M. S., Midtiby, H. S., Kruger, N., and Jorgensen, R. N. (2014). Statistics-based segmentation using a continuous-scale naive bayes approach. *Comput. Electron. Agric.* 109, 271–277. doi: 10.1016/j.compag.2014.10.009

Li, Z. Q., Chen, L. Q., Zheng, Q., Dou, X. Y., and Yang, L. (2019). Control of a path following caterpillar robot based on a sliding mode variable structure algorithm. *Biosyst. Eng.* 186, 293–306. doi: 10.1016/j.biosystemseng.2019.07.004

Liu, H. J., Sun, H., Li, M. Z., and Iida, M. (2020). Application of color featuring and deep learning in maize plant detection. *Remote Sens.* 12, 2229. doi: 10.3390/rs12142229

Liu, X. E., Li, M. J., Sun, Y. L., and Deng, X. Y. (2010). Support vector data description for weed/corn image recognition. *J. Food Agric. Environ.* 8, 214–219. doi: 10.1023/B:MACH.0000008084.60811.49

Luo, X., Qu, H., Wang, Y., Yi, Z., Zhang, J., and Zhang, M. (2022). Supervised learning in multilayer spiking neural networks with spike temporal error backpropagation. *IEEE Trans. Neural Netw. Learn. Syst.* doi: 10.1109/TNNLS.2022.3164930

Ma, Z. H., Tao, Z. Y., Du, X. Q., Yu, Y. X., and Wu, C. Y. (2021). Automatic detection of crop root rows in paddy fields based on straight-line clustering algorithm and supervised learning method. *Biosyst. Eng.* 211, 63–76. doi: 10.1016/j.biosystemseng.2021.08.030

Manavalan, R. (2020). Automatic identification of diseases in grains crops through computational approaches: a review. *Comput. Electron. Agric.* 178, 5802. doi: 10.1016/j.compag.2020.105802

Mavridou, E., Vrochidou, E., Papakostas, G. A., Pachidis, T., and Kaburlasos, V. G. (2019). Machine vision systems in precision agriculture for crop farming. *J. Imaging* 5, 89. doi: 10.3390/jimaging5120089

Montalvo, M., Pajares, G., Guerrero, J. M., Romeo, J., Guijarro, M., Ribeiro, A., et al. (2012). Automatic detection of crop rows in maize fields with high weeds pressure. *Expert Syst. Appl.* 39, 11889–11897. doi: 10.1016/j.eswa.2012.02.117

Ospina, R., and Noguchi, N. (2019). Simultaneous mapping and crop row detection by fusing data from wide angle and telephoto images. *Comput. Electron. Agric.* 162, 602–612. doi: 10.1016/j.compag.2019.05.010

Otsu, N. (2007). A threshold selection method from gray-level histograms. *IEEE Trans. Syst. Man Cybern.* 9, 62–66. doi: 10.1109/TSMC.1979.4310076

Pan, Z., Zhang, M., Wu, J., Wang, J., and Li, H. (2021). Multi-tone phase coding of interaural time difference for sound source localization with spiking neural networks. *IEEE/ACM Trans. Audio Speech Lang. Process.* 29, 2656–2670. doi: 10.1109/TASLP.2021.3100684

Ponnambalam, V. R., Bakken, M., Moore, R. J. D., Gjevestad, J. G. O., and From, P. J. (2020). Autonomous crop row guidance using adaptive multi-roi in strawberry fields. *Sensors* 20, 5249. doi: 10.3390/s20185249

Rabab, S., Badenhorst, P., Chen, Y. P. P., and Daetwyler, H. D. (2021). A template-free machine vision-based crop row detection algorithm. *Precision Agric.* 22, 124–153. doi: 10.1007/s11119-020-09732-4

Radcliffe, J., Cox, J., and Bulanon, D. M. (2018). Machine vision for orchard navigation. *Comput. Ind.* 98, 165–171. doi: 10.1016/j.compind.2018.03.008

Rovira-Mas, F., Saiz-Rubio, V., and Cuenca-Cuenca, A. (2021). Augmented perception for agricultural robots navigation. *IEEE Sens. J.* 21, 11712–11727. doi: 10.1109/JSEN.2020.3016081

Roy, K., Jaiswal, A., and Panda, P. (2019). Towards spike-based machine intelligence with neuromorphic computing. *Nature* 575, 607–617. doi: 10.1038/s41586-019-1677-2

Shrivastava, V. K., and Pradhan, M. K. (2021). Rice plant disease classification using color features: a machine learning paradigm. *J. Plant Pathol.* 103, 17–26. doi: 10.1007/s42161-020-00683-3

Sui, B., Jiang, T., Zhang, Z., and Pan, X. (2020). Ecgan: An improved conditional generative adversarial network with edge detection to augment limited training data for the classification of remote sensing images with high spatial resolution. *IEEE J. Select. Top. Appl. Earth Observat. Remote Sens.* 14, 1311–1325. doi: 10.1109/JSTARS.2020.3033529

Tang, J. L., Chen, X. Q., Miao, R. H., and Wang, D. (2016). Weed detection using image processing under different illumination for site-specific areas spraying. *Comput. Electron. Agric.* 122, 103–111. doi: 10.1016/j.compag.2015.12.016

Tenhunen, H., Pahikkala, T., Nevalainen, O., Teuhola, J., Mattila, H., and Tyystjarvi, E. (2019). Automatic detection of cereal rows by means of pattern recognition techniques. *Comput. Electron. Agric.* 162, 677–688. doi: 10.1016/j.compag.2019.05.002

Vishnoi, V. K., Kumar, K., and Kumar, B. (2022). A comprehensive study of feature extraction techniques for plant leaf disease detection. *Multimed Tools Appl.* 81, 367–419. doi: 10.1007/s11042-021-11375-0

Vrochidou, E., Oustadakis, D., Kefalas, A., and Papakostas, G. A. (2022). Computer vision in self-steering tractors. *Machines* 10, 129. doi: 10.3390/machines10020129

Wang, Q. Q., Jiang, Y. C., Li, L. H., Qin, J. W., and Chen, L. Q. (2021). Performance analysis of a spring-tooth drum pickup of straw baler via coupling simulation. *Int. J. Agric. Biol. Eng.* 14, 159–165. doi: 10.25165/j.ijabe.20211404.6576

Winterhalter, W., Fleckenstein, F. V., Dornhege, C., and Burgard, W. (2018). Crop row detection on tiny plants with the pattern hough transform. *IEEE Rob. Automat. Lett.* 3, 3394–3401. doi: 10.1109/LRA.2018.2852841

Xue, Y., Tang, T., Pang, W., and Liu, A. X. (2020). Self-adaptive parameter and strategy based particle swarm optimization for large-scale feature selection problems with multiple classifiers. *Appl. Soft Comput.* 88, 106031. doi: 10.1016/j.asoc.2019.106031

Xue, Y., Tang, Y., Xu, X., Liang, J., and Neri, F. (2021a). Multi-objective feature selection with missing data in classification. *IEEE Trans. Emerg. Top. Comput. Intell.*, 6, 355–364. doi: 10.1109/TETCI.2021.3074147

Xue, Y., Zhu, H., Liang, J., and Słowik, A. (2021b). Adaptive crossover operator based multi-objective binary genetic algorithm for feature selection in classification. *Knowl. Based Syst.* 227, 107218. doi: 10.1016/j.knosys.2021.107218

Yu, Y., Bao, Y. D., Wang, J. C., Chu, H. J., Zhao, N., He, Y., et al. (2021). Crop row segmentation and detection in paddy fields based on treble-classification otsu and double-dimensional clustering method. *Remote Sens.* 13, 901. doi: 10.3390/rs13050901

Zhang, M., Qu, H., Belatreche, A., Chen, Y., and Yi, Z. (2018a). A highly effective and robust membrane potential-driven supervised learning method for spiking neurons. *IEEE Trans. Neural Netw. Learn. Syst.* 30, 123–137. doi: 10.1109/TNNLS.2018.2833077

Zhang, M., Wang, J., Wu, J., Belatreche, A., Amornpaisannon, B., Zhang, Z., et al. (2021). Rectified linear postsynaptic potential function for backpropagation in deep spiking neural networks. *IEEE Trans. Neural Netw. Learn. Syst.* 33, 1947–1958. doi: 10.1109/TNNLS.2021.3110991

Zhang, X. Y., Li, X. N., Zhang, B. H., Zhou, J., Tian, G. Z., Xiong, Y. J., et al. (2018b). Automated robust crop-row detection in maize fields based on position clustering algorithm and shortest path method. *Comput. Electron. Agric.* 154, 165–175. doi: 10.1016/j.compag.2018.09.014

Zhou, Y., Yang, Y., Zhang, B. L., Wen, X., Yue, X., and Chen, L. Q. (2021). Autonomous detection of crop rows based on adaptive multi-roi in maize fields. *Int. J. Agric. Biol. Eng.* 14, 217–225. doi: 10.25165/j.ijabe.20211404.6315

Conflict of Interest: DL was employed by JD AI Research.

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

Publisher's Note: 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: 11 July 2022
doi: 10.3389/fpls.2022.915506

Effects of Different Environment-Friendly Gibberellic Acid Microcapsules on Herbicide Injury of Wheat

*Guisen Zhang[1], Tao Ma[2], Yong Cheng[3], Jianing Wang[1], Lang Liu[1] and Baojun Zhang[1]**

[1] College of Plant Protection, Shanxi Agricultural University, Jinzhong, China, [2] Shanxi Inspection and Testing Center (Shanxi Institute of Standards and Metrology), Taiyuan, China, [3] Bayer Crop Science (China) Co., Ltd., Beijing, China

Environmentally friendly microcapsules are becoming more and more widely used due to the increasing demand for environmental safety in pesticides. To compare the impact of differences in wheat herbicide phytotoxicity of different gibberellic acid microcapsule suspensions, two microcapsule suspensions were separately formulated using the phase transfer method and the *in situ* polymerization method, and the key performance indicators are the size of microcapsules and the degree of encapsulation. Meanwhile, through field trials, the pharmacological and detoxification effects of different types of microcapsule suspensions on the herbicide methyldisulfuron in wheat fields were compared. The microcapsule suspension was prepared by the phase transfer method and the particle sizes D_{10}, D_{50}, and D_{90} are 0.990, 2.136, and 5.201 μm, respectively; the microcapsule suspension was prepared by the *in situ* polymerization, the particle sizes D_{10}, D_{50}, and D_{90} are 4.365, 8.547, and 16.782 μm, respectively. The encapsulation rate of the microcapsules prepared by the phase transfer method and the *in situ* polymerization method was 86.9% and 91.2%, being higher than 80%, the national standard for capsules. Meanwhile, the release rate conforms to first-order release kinetics in 0–4 days and zero-order release kinetics in 5–28 days. The plot trials' result showed that the detoxification effect of the microcapsules prepared by the *in situ* polymerization method was significantly better than the detoxification effect of the microcapsules prepared by the phase transfer method and the control agent. The growth index of wheat was higher than that of the untreated check after using the agent.

Keywords: microcapsules, gibberellic acid, brassinolide, phytotoxicity, soluble powder

OPEN ACCESS

Edited by:
Jian Su,
Nanjing University of Information Science and Technology, China

Reviewed by:
Zhixin Zhou,
Hangzhou Dianzi University, China
Xinyan Wang,
Xi'an Peihua University, China
Baoqing Zhang,
Guangxi Academy of Agricultural Science, China
Jie Wang,
Tobacco Research Institute (CAAS), China

***Correspondence:**
Baojun Zhang
sxhyph@163.com

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

Received: *08 April 2022*
Accepted: *23 May 2022*
Published: *11 July 2022*

Citation:
Zhang G, Ma T, Cheng Y, Wang J, Liu L and Zhang B (2022) Effects of Different Environment-Friendly Gibberellic Acid Microcapsules on Herbicide Injury of Wheat. Front. Plant Sci. 13:915506. doi: 10.3389/fpls.2022.915506

INTRODUCTION

Microcapsule technology is widely used in global agricultural production because microcapsules are more environmentally friendly compared with other pesticide formulations, such as pesticide emulsion, wettable powder, and microemulsion. At present, inorganic or organic polymer materials are mainly coated on the surface of the active ingredient solution *via* physical and chemical methods to form solid particles with a certain shape. The diameter of such solid particles is generally between 10 and 1,000 microns (Cui et al., 2018), and the release of active ingredients occurs mainly through two mechanisms: osmotic diffusion and capsular rupture (Wang, 2009). Therefore, microcapsule

technology can provide controlled release, prolong residual efficacy, and improve products' storage stability, and all the features can provide a significant value to pesticides applied in agricultural production, which has a high application value (Liu et al., 2018).

Gibberellic acid is a plant growth regulating hormone with good performance and widely used in global agricultural production. Gibberellic acid can provide excellent performance in promoting plant growth and cell differentiation, delaying senescence, and increasing crop tolerance. Gibberellic acid applied during the wheat growth stage promotes early wheat maturity, improves wheat quality and yield, and at the same time can improve wheat crop tolerance (Liu et al., 2015; Chen, 2018; Yu et al., 2019). However, the performance of gibberellic acid in the field is easily affected due to high temperature, moisture, and other environmental factors, sometimes, the efficacy in the field is reduced due to the decomposition of an active ingredient. Microcapsule technology can prolong residual efficacy *via* reducing the decomposition of an active ingredient. Methylsulfuron is a sulfonylurea herbicide, which can control a variety of grass weeds and is widely used in wheat fields globally (Zhou et al., 2019). However, if it is used improperly, such as repeated spraying and multiple spraying, or under unfavorable environmental conditions, such as rain, cooling, and flooding, severe pesticide phytotoxicity may occur after application and cause yield losses in the field, and urea or plant growth regulators (such as gibberellic acid) applied in time will reduce the impact on crops and yield (Cao et al., 2018). However, there are no reports on the microencapsulation of gibberellic acids and their detoxification effect on herbicide phytotoxicity. Microcapsules as environmentally friendly pesticide formulations play an important role. The author used the phase transfer method and *in situ* polymerization to prepare microcapsule suspensions of different types of gibberellic acid, and applied the prepared microcapsule suspensions to the field. The result showed that microencapsulation of gibberellic acid can reduce wheat herbicide phytotoxicity caused by methyl sulfuron, and this research provides a theoretical basis and data support for subsequent practical agricultural production practice in the future.

RELATED WORKS

Main Instruments and Reagents

Materials

Sodium hydroxide, 36% concentrated hydrochloric acid, urea, formaldehyde, ethyl acetate, ethyl cellulose, butyl phthalate, polyvinyl alcohol, n-dodecane, methyl oleate, xanthan gum, and ethylene glycol are all analytical grade. The above reagents were purchased from Sinopharm Chemical Reagent Co., Ltd. Tween-80 and Tween-20 are chemically pure and purchased from Hai'an Petrochemical Plant in Jiangsu Province; 30 g·L^{-1} methyldisulfuron-dispersible oil suspension agent (Shandong Bio-Biotechnology Co., Ltd.), 10% Gibberellic acid soluble powder (Zhejiang Qianjiang Biochemical Co., Ltd.), 90.3% Gibberellic acid (GA3) original medicine was provided by Sichuan Guoguang Agrochemical Co., Ltd.

Instruments

Instruments include JT2003B electronic balance (Yuyao Jinnuo Balance Instrument Co., Ltd.), E-100 optical microscope (Nikon Corporation), GZX-GF101 electric blast drying oven (Shanghai Yuejin Medical Instrument Co., Ltd.), WJL-602 laser particle size analyzer (Shanghai Yidian Physical and Optical Instrument Co., Ltd.), HH-2 Electric Heating Constant Temperature Water Bath (Beijing Kewei Yongxing Instrument Co., Ltd.), GC-4000 A Liquid Chromatograph (Beijing Dongxi Analytical Instrument Co., Ltd.), JP-100 Ultrasonic Cell Crusher (Shenzhen Jiemeng Cleaning Equipment Co., Ltd.), and Digital Vernier Calipers (Shijiazhuang Woma Tools Co., Ltd.).

Preparation of 10% Gibberellic Acid Microcapsule Suspension

Preparation of Microcapsule Suspensions by the Phase Transfer Method

Aqueous Phase Preparation

Approximately 40 ml of an aqueous solution containing 0.5 g of Tween-80 and 1.5 g of PVA was prepared.

Organic Phase Preparation

Approximately 1 g of ethyl cellulose and 2 g of methyl oleate were weighed into a beaker, added with 5.54 g of gibberellic acid and 10 ml of ethyl acetate solution, and then stirred to completely dissolve ethyl cellulose and gibberellic acid.

Microcapsule Preparation

The prepared aqueous solution was poured into a three-necked flask equipped with a mechanical stirrer by adjusting the rotation speed to 1,500 r·min^{-1}, controlling the water temperature at 30°C, and adding the organic phase dropwise at a speed of 3 ml·min^{-1}. After dropping, stirring was continued for 20 min by slowly raising the temperature to 60°C and adjusting the rotation speed to 550 r·min^{-1}, and when the ethyl acetate is evaporated, it was made up with 50 g water to obtain 10% gibberellic acid microcapsule suspension.

Preparation of Microcapsule Suspensions by *in situ* Polymerization

Preparation of the Oil Phase

About 6.18 g of gibberellic acid was weighed, dissolved in 12.0 g of ethyl acetate, and used as the oil phase.

Preparation of the Water Phase

In a three-neck flask equipped with a thermometer and stirring device, formaldehyde and urea were mixed according to the molar ratio of 1.75, dissolved and added with an appropriate amount of deionized water, and then added to the three-neck flask and stirred at a rate of 200 r·min^{-1}. Meanwhile the pH was adjusted to 8.0 with 0.5 mol·L^{-1} sodium hydroxide solution and then warmed up to 70°C at a rate of 2°C·min^{-1}, and the reaction was carried out for 1 h after warming up to obtain formaldehyde–urea prepolymer solution (Zhang, 2012). A further 8.0 g of 30% urea–formaldehyde resin prepolymer was mixed with 2.0 g of emulsifier T-20 and 21.82 g of water as the aqueous phase.

Preparation of Microcapsules

The oil and water phases were mixed, sheared for 60 s at 10,000 $r \cdot min^{-1}$, transferred to a 100-ml three-necked flask, stirred at 30°C and 800 $r \cdot min^{-1}$, and hydrochloric acid was used (1%). The pH was adjusted to 4.5, heated to 60°C, and stirred for 60 min. Approximately 0.1 g of xanthan gum and 2.0 g of ethylene glycol were added and stirring was continued for 30 min. Sodium hydroxide (20%) was used to adjust the pH to 9, and finally water was replenished to 50 g to obtain a 10% gibberellic acid microcapsule suspension agent with urea–formaldehyde resin as a wall material (Zhang, 2011).

Characterization of 10% Gibberellic Acid Microcapsule Suspension

Determination of the Appearance and Particle Size of Microcapsule Suspension

An optical microscope was used to change the four fields of view, observe the morphology of the microcapsules, and, in the meantime, measure the average particle size of the microcapsules with a laser particle size distribution analyzer.

Determination of Encapsulation Efficiency

Approximately 0.05 g of the microcapsule suspension sample was accurately weighed (accurately to 0.0001 g), and an appropriate amount of methanol was added, the capsule was broken with an ultrasonic cell pulverizer, and then made up to 100 ml volume with methanol. The content of active ingredients in the microcapsule suspension was determined by liquid chromatography (Zhang, 2012).

Calculation of microencapsulation efficiency:

$$\text{Encapsulation rate (\%)} = (\text{microcapsule core material content} / \text{core material content in raw materials}) \times 100. \quad (1)$$

Operating conditions of liquid chromatography: column temperature: room temperature; detection wavelength: 210 nm; mobile phase: methanol + water + formic acid = 40 + 60 + 0.05 (V/V); flow rate: 0.8 $ml \cdot min^{-1}$; injection volume: 20 μl (Ye et al., 1996).

Study on the Release Kinetics of Microcapsules in Organic Solvents

The prepared microcapsule suspension was filtered and dried for use. A 0.1000 g of dry sample was weighed, transferred to a 100-ml volumetric flask to break the capsule with an ultrasonic cell disruptor, and then made up to 100 ml with xylene. Liquid chromatography was used to determine the drug loading in the dried microcapsule sample.

About 0.2000 g of the sample was weighed again, transferred to a three-necked flask containing 200 ml of xylene, and stirred at 40°C and 400 $r \cdot min^{-1}$. Samples were taken once a day for the first 5 days, and samples were taken every 2 days from the 6th day. Aspirate 5 ml of liquid while adding 5 ml of xylene. The extracted liquid was filtered to determine the cumulative release amount at each point by liquid chromatography, and the cumulative release percentage (%) was calculated. The formula was calculated according to formula 1: (Zhang et al., 2012).

Gibberellic Acid Microcapsule Suspension Herbicide Detoxification

Test Process

The wheat variety Jinmai 101 provided by the Wheat Research Institute of Shanxi Academy of Agricultural Sciences was tested. It is a semi-winter variety with a growth period of 235 days and an average plant height of about 88 cm.

The trial was conducted in Sangzi village, Hucun town, Taigu County, Jinzhong City, Shanxi province in 2018–2019. The trial field conditions are flat field, sandy loam soil, 1.57% organic, and pH 7.4. For test field preparation, the field was irrigated at 5 days before sowing to ensure good field conditions for mechanical planting. And, the wheat was sown at 24 September 2018, and the seeding amount is 300 $kg \cdot hm^{-2}$. The test was planted in a north-south direction *via* a randomized block design, and the plot size is 25 m^2 (length 5 m, width 5 m, 13 rows per plot, 0.38 m pitch, and 0.50 m plot spacing), and 1.00 m width walkway between replications, and even crop vigor among four replications. After returning to the green (21 March 2019), methyldisulfuron was applied at 30 $g \cdot L^{-1}$ and the dosage was 30 g $ai \cdot ha^{-1}$. Following the finding of methyldisulfuron phytotoxicity symptom, such as yellowing of leaves or the weakness of plants or reduced plant height or stunted growth (26 March 2019) on wheat, gibberellic acid microcapsule suspension was applied to wheat with four replications. The treatment list is presented in **Table 1**.

Survey of the Results

Samples were collected before and at 21 days after the application of the microcapsule suspension, five points per plot, and five plants per point. The height of each plant was investigated, and the plant height growth and growth promotion rate were calculated. The calculation formula was as follows: formulas (2) and (3). Meanwhile, samples were collected before harvest once again, five points per plot, and five plant per point. Ear length, ear diameter, single ear weight, and 1,000-kernel weight were investigated, and the yield of each plot was assessed.

$$\text{Plant height growth} = \text{plant height (cm) after medicine} - \text{plant height (cm) before medicine.} \quad (2)$$

Data Analysis

Variance analysis on the above mentioned index traits was analyzed by the SPSS data processing system, and multiple comparisons were performed to analyze the differences between the different agents *via* a new complex range method.

TABLE 1 | Concentration of the reagents used in the test field.

No.	Pharmacy	Active ingredient dosage (ml·ha^{-1})	Remarks
1	10% gibberellic acid microcapsule suspension (Microcapsule 1)	250	Prepared by the phase transfer method
2		500	
3		750	
4	10% gibberellic acid microcapsule suspension (Microcapsule 2)	250	Prepared by *in situ* polymerization
5		500	
6		750	
7	10% gibberellic acid soluble powder	250	—
8		500	
9		750	
10	Untreated check	–	-

Microcapsule 1 represents the formulation prepared by the phase transfer method and microcapsule 2 represents the formulation prepared by the in situ polymerization method, and so on.

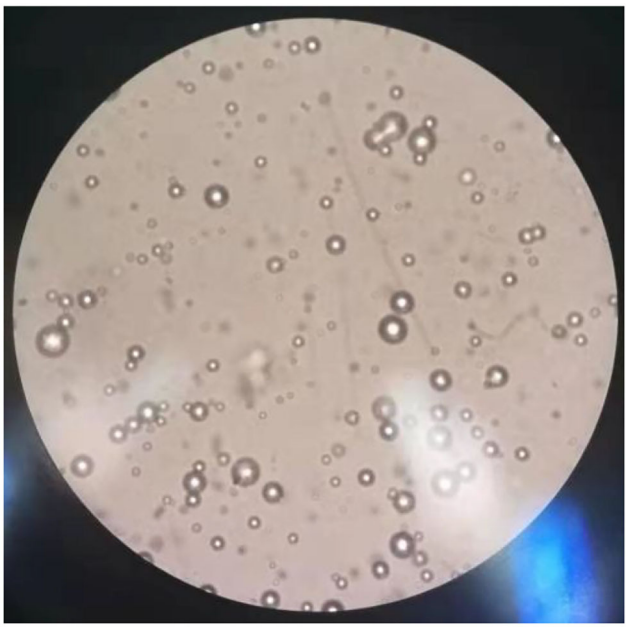


FIGURE 1 | Microscopic images on microcapsule particle size prepared by the phase transfer method (100×).

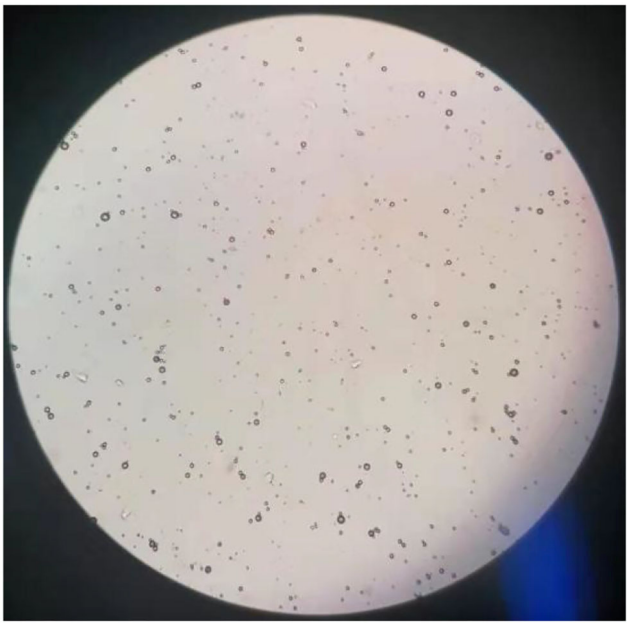


FIGURE 2 | Microscopic images on microcapsule particle size prepared by *in situ* polymerization (100×).

RESULTS AND DISCUSSION

Determination of Performance Indicators of Microcapsule Suspensions

Determination of the Appearance and Particle Size of Microcapsule Suspensions

The results of observations of microcapsules 1 and 2 using an optical microscope are shown in **Figures 1**, **2**. The particle size measured by the laser particle size distribution analyzer showed that: for microcapsule 1, D_3 is 0.697 μm, D_{10} is 0.990 μm, D_{50} is 2.136 μm, D_{90} is 5.201 μm, and D_{97} is 7.987 μm, $\delta = 1.971$; for microcapsule 2, D_3 is 3.241 μm, D_{10} is 4.365 μm, D_{50} is 8.547 μm, D_{90} is 16.782 μm, and D_{97} is 22.946 μm, $\delta = 1.453$; particle size. Distribution diagrams are shown in **Figures 3**, **4**. Comprehensively, the result showed that microcapsule 1 had a smaller particle size, worse monodispersity, and wider particle size distribution compared with microcapsule 2. This phenomenon occurs during the microcapsule preparation process. The selection of the emulsifier is related to the selection of a suitable emulsifier when preparing the aqueous phase, which can emulsify the oil phase into fine oil droplets and disperse uniformly in the aqueous phase.

Determination of Encapsulation Efficiency

The encapsulation rate of microcapsules 1 and 2 was 86.9% and 91.2%. The results showed that the encapsulation rate of

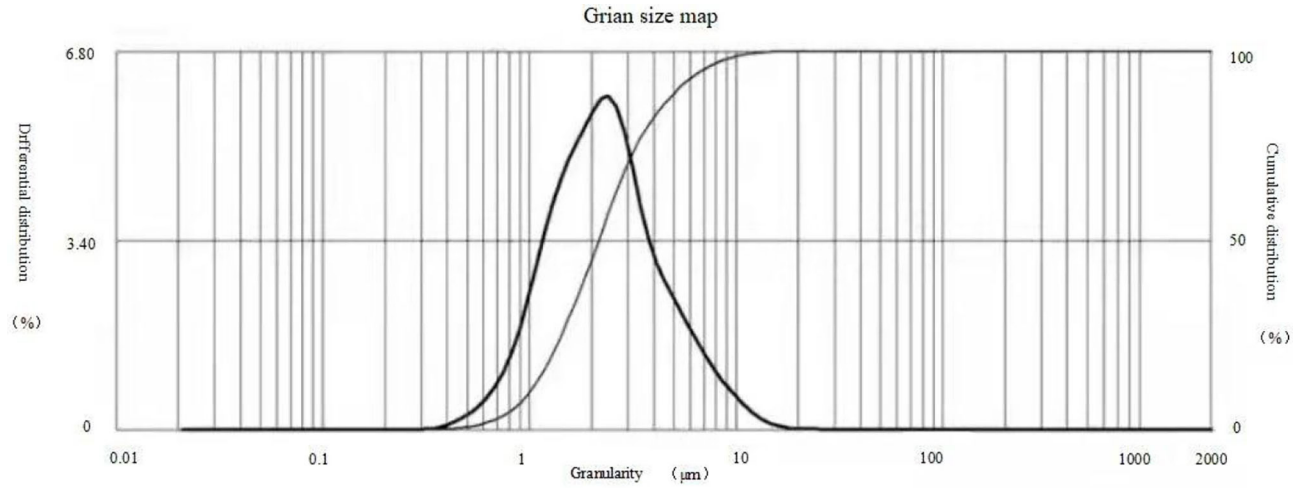


FIGURE 3 | Particle size distribution of microcapsule 1.

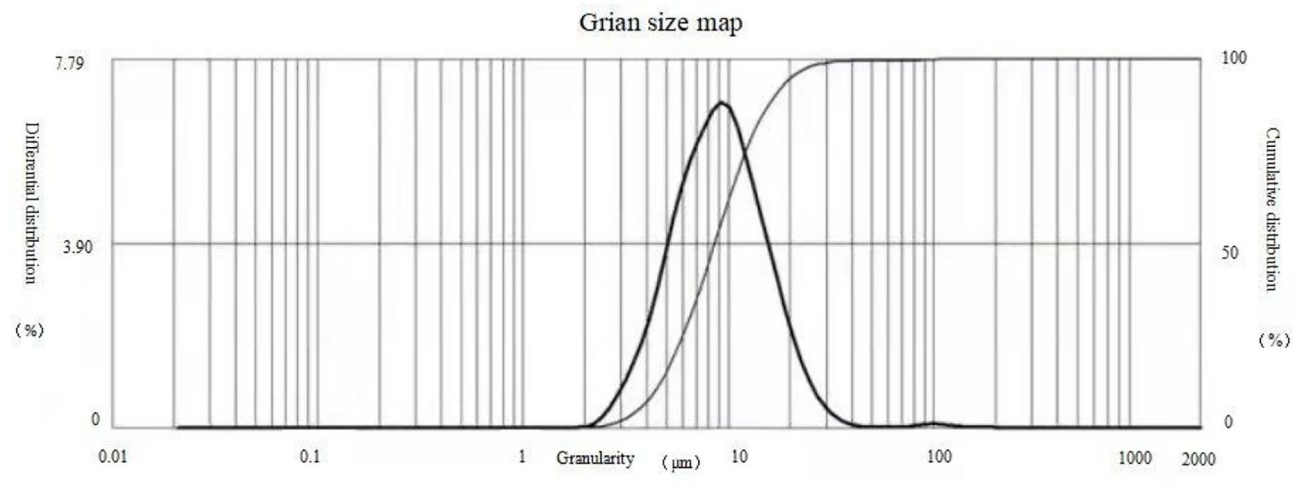


FIGURE 4 | Particle size distribution of microcapsule 2.

microcapsules prepared by the *in situ* polymerization method is high. It is possible that microcapsule 2 is formed by encapsulating the active ingredients after polymerization of the capsule wall monomer, the formed capsule is denser; while microcapsule 1 is a capsule prepared by the physical method, the phase transfer occurs in the capsule wall. Due to the influence of the volatilization rate of organic solvents, the conditions are difficult to grasp, and the cysts formed are not uniform.

Study on the Release Kinetics of Microcapsules Into Organic Solvents

The release profiles of the microcapsules prepared by the two methods are shown in **Figures 5**, **6**. For microcapsule 1, the release rate on the 1st day was 13.4%, and the cumulative release rate in the first 5 days was 32.5%, indicating that the microcapsules had a rapid release process in the early stage. From 0 to 5 days, by fitting the kinetic equation with time and cumulative release, the zero-order release kinetic equation of microcapsule 1 is $y = 5.6\times + 10.45$ ($R^2 = 0.8659$, Q is the cumulative release percentage), and the first-order release kinetic theoretical equation is $\mathrm{Ln}(1 - Q) = -0.0725\times + 4.5022$ ($R^2 = 0.8828$). The R^2 value of the zero-order release kinetic equation is lower than that of the first-order release kinetic equation. This means that the release of microcapsule 1 complies with the first-order kinetic equation from 0 to 4 days, and the release behavior of the active ingredients in the microcapsules has entered a stable phase from the 5th day. From 5 to 28 days, the zero-order release kinetic equation is $y = 2.5519\times + 20.792$ ($R^2 = 0.989$), and the first-order release kinetic equation is $\mathrm{Ln}(1 - Q) = -0.0792\times + 4.7626$ ($R^2 = 0.9581$). The R^2 value of the zero-order release kinetic equation is greater than that of the first-order release kinetic equation, which means that the release of

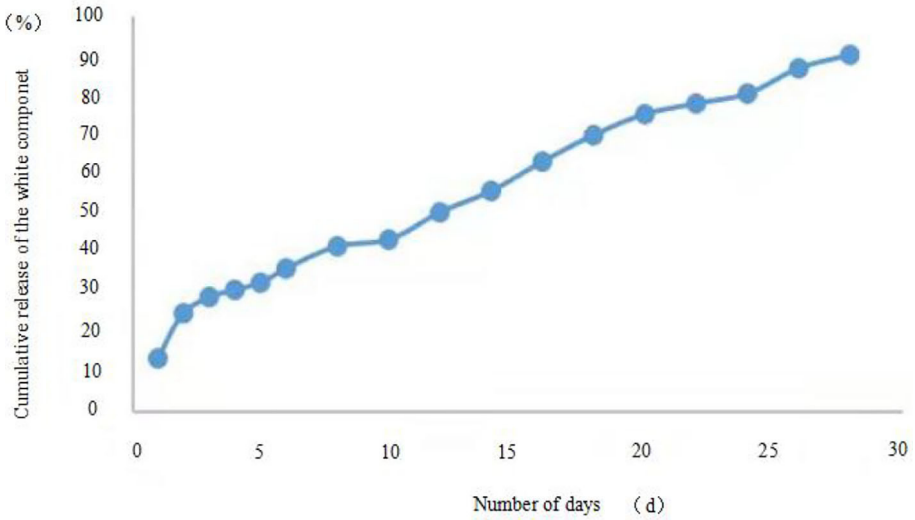


FIGURE 5 | Microcapsule 1 release curve.

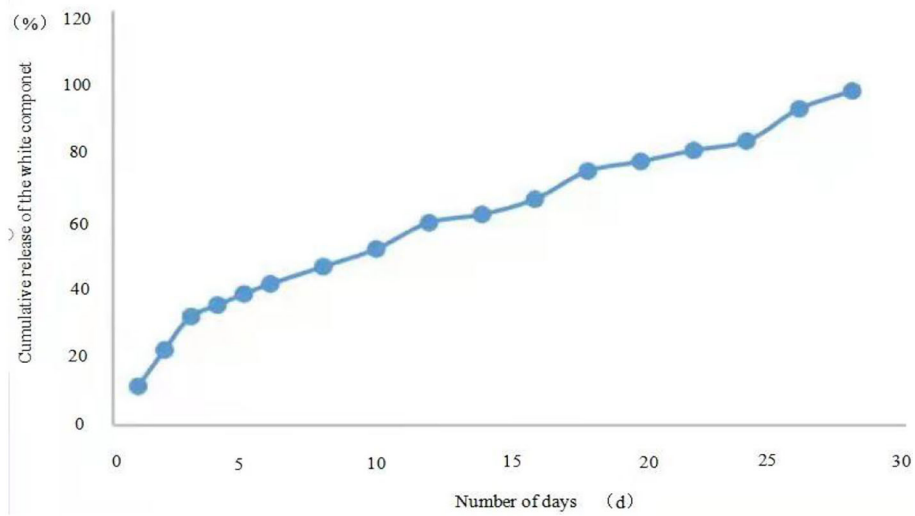


FIGURE 6 | Microcapsule 2 release curve.

microcapsule 1 fits the zero-order kinetic equation from 5 to 28 days.

Similarly, the first-order release kinetic equation of microcapsule 2 released in 0–4 days is Ln $(1–Q) = -0.1054\times + 4.5772$ ($R^2 = 0.9676$), and the zero-order kinetic equation is $y = 8.02\times + 4.7$ ($R^2 = 0.9568$), the release from 5 to 28 days the first-order kinetic equation is Ln$(1–Q) = -0.0998\times + 4.8952$ ($R^2 = 0.8367$), the zero-order kinetic equation is $y = 2.4675\times + 26.422$ ($R^2 = 0.993$), which means that the release of microcapsule 2 from 0 to 4 days conforms to the first-order kinetic equation, and the release of microcapsules in 5–28 days conforms to the zero-order kinetic equation.

Research on Detoxification

Effect of Pesticides on Wheat Growth

The effects of the different pesticides on wheat growth are shown in **Table 2**. At 14 days after application, plant height in all gibberellic acid treatments is significantly higher than that in untreated checks, and a clear dosage response was found in different gibberellic acid treatments. Meanwhile, there are no significant differences between microcapsules 1 and 2 at the same dosage, which are significantly lower compared with gibberellic acid soluble powder treatment at the same dosage. At 28 days after application, plant height in all gibberellic acid treatments is still significantly higher than that in untreated checks. Meanwhile, plant height in microcapsule treatments

TABLE 2 | Detoxification effects of different pesticides on wheat.

Treatments	Amount of active ingredients (ml·ha^{-1})	Growth height (cm)		Growth promotion rate	
		14 days after application	28 days after application	14 days after application	28 days after application
Microcapsule 1	250	10.28d	31.96e	11.14d	19.50e
	500	13.48c	33.93cd	45.77c	26.88cd
	750	15.00b	35.37b	62.18b	32.26b
Microcapsule 2	250	10.70d	31.18e	15.71d	16.58e
	500	13.48c	34.71bc	45.77c	29.80bc
	750	15.24b	37.78a	64.79b	41.26a
10% gibberellic acid soluble powder	250	12.96c	28.42f	40.13c	6.26f
	500	14.80b	31.07e	59.99b	16.17e
	750	18.65a	33.42d	101.70a	24.97d
Untreated Check	–	9.25e	26.74 g		

The meaning of the letters a–d represents the significance of the difference, and the different letters indicate that the results are significantly different from each other.

TABLE 3 | An effect of the different pesticides on wheat yield.

Elixir	Amount of active ingredient (ml ·ha^{-1})	Ear diameter (cm)	Ear length (cm)	Single spike weight (g)	1,000 kernel weight (g)	Yield (kg·hm^{-2})	Yield increase (%)
Microcapsule 1	250	1.18ef	7.51d	2.22cde	40.01e	6893.45ef	3.45ef
	500	1.23cd	7.72bc	2.39b	43.23cd	7463.73bcd	12.01bcd
	750	1.30ab	7.89ab	2.59a	45.71ab	8624.31a	29.43a
Microcapsule 2	250	1.21de	7.53d	2.24cd	41.71de	7073.54def	6.16def
	500	1.25bc	7.78ab	2.38b	44.10bc	7763.88b	16.52b
	750	1.32a	7.94a	2.67a	46.75a	8824.41a	32.43a
10% gibberellic acid soluble powder	250	1.15f	7.34e	2.19de	37.52f	6683.34f	0.30f
	500	1.21de	7.60cd	2.31bc	39.83e	7243.62cde	8.71cde
	750	1.26bc	7.75bc	2.40b	41.57de	7623.81bc	14.41bc
Blank control	–	1.08 g	7.18f	2.12e	35.25 g	6663.33f	–

Different lowercase letters after the same column of data indicate a significant difference of 5%.

is significantly higher than that in gibberellic acid soluble powder treatment at the same dosage due to the longer residual control from the microcapsule compared with the soluble powder formulation, in terms of comparison of two types of microcapsule, plant height in microcapsule 2 @ 750 ml/ha is significantly higher than microcapsule 1 @ 750 ml/ha, and plant height in microcapsules 1 and 2 @ 250–500 ml/ha is comparable to each other at the same dosage.

Effect of Pesticides on Wheat Yield

The effects of spraying different pesticides on wheat yield are shown in **Table 3**. A clear dosage response was found in all the assessment factors for all the three gibberellic acid formulations. Yield performance from three gibberellic acid formulations @ 250 ml/kg is comparable with the untreated check due to comparable or slightly better performance in the assessed factors, such as ear diameter, ear length, single ear weight, and 1,000-grain weight, from different gibberellic acid formulations compared with the untreated check, and a significantly higher yield from three gibberellic acid formulations @ 500–750 ml/kg with the untreated check due to significantly better performance in the assessed factors, such as ear diameter, ear length, single ear weight, and 1,000-grain weight. Meanwhile, the yield of both microcapsules @ 750 ml/ha is significantly higher than gibberellic acid soluble powder @ 750 ml/ha, which is driven by 1,000-grain weight and single ear weight, and all the assessed factors of microcapsules 2 @ 750 ml/ha are significantly better than gibberellic acid soluble powder @ 750 ml/ha, and the assessed factors of microcapsule 1 @ 750 ml/ha are better than or comparable with gibberellic acid soluble powder @ 750 ml/ha. Comprehensively, microcapsule 2 provided the best effect in the field especially on yield, followed by microcapsule 1, which is better than the control agent. This shows that the microcapsules can significantly increase the yield of wheat through yield loss reduction due to the damage caused by pesticides (Wang et al., 2000).

DISCUSSION

Characterization of Pesticide Microcapsule Release Properties

Microcapsule suspensions, as a new type of controlled released formulation, have the outstanding advantages of prolonging the residual efficacy from the active ingredient through slow release (Liu and Xie, 2005). Release is an important indicator, which can directly reflect the effect of sustained release and durability of microcapsules, and it is of great significance for evaluating and using pesticide microcapsules (Zhao et al., 2007). Meanwhile, the release of the microcapsules has a significant impact by means of environmental conditions, such as temperature, pH, and moisture. And, the release mechanism is usually done through physical routes, such as dissolution, diffusion, penetration, and ion exchange, and chemical pathways, such as active ingredients or enzyme degradation (Chen, 2015; Xue, 2019), normally, there are two phases in the release process, which are rapid release and slow release phase. In the rapid release phase, an active component on the outer wall of the capsule may not be coated, resulting in a large amount of initial release, which is consistent with the "burst release effect" of the initial release of microcapsules (Jegat and Taverdet, 2000). In the slow-release phase, the active ingredient diffuses through the wall of the capsule over time, the release is gentle and the cumulative release amount increases slowly; through these two phases, the microcapsules can be guaranteed to have a rapid and sustained effect. During the experiment, the microcapsules were placed in an organic solvent to simulate the actual release of the microcapsules through diffusion. There are two stages in the release process of microcapsules 1 and 2, which meet the characteristics of slow release of microcapsules. The mechanism and characterization of the release properties of pesticide microcapsules prepared from different capsule wall materials need to be further explored.

Comparison of Field Effects of Different Methods for Preparing Microcapsules

This experiment verifies that the microcapsules prepared by the two different methods have different effects in their actual use. The specific performance is that the microcapsules prepared by the *in situ* polymerization method are more effective than the microcapsules prepared by the phase transfer method, both of which are higher than soluble powder. The better performance should be caused by a smaller particle size of microcapsule 2 compared with microcapsule 1. The smaller particle size means that the microcapsule is easier to attach to the surface of the crop and not easy to fall. Meanwhile, the emulsifier and processing technology are also related; on the other hand, the release is faster, and the amount released after 28 days is higher in microcapsule 2 than in microcapsule 1. Through the release of microcapsules, gibberellic acid can be continuously released into the environment, to extend the efficacy in promoting crop growth, and continuously increase crop tolerance against herbicide damage. However, the efficacy at 14 days is not as good as the soluble powder formulation, which may be caused by the microcapsules that have not yet been completely released. The control agent has a higher amount of gibberellic acid active ingredients on the surface of wheat than the microcapsules, at 28 days, the efficacy is significantly better than the soluble powder formulation, which demonstrates the longer durability of the microcapsules and the advantages of resisting the effects of the natural environment on the active ingredients.

CONCLUSION

Two gibberellic acid microcapsule suspensions prepared by the phase transfer method and the *in situ* polymerization method have a uniform particle size distribution. The encapsulation efficiency and release characteristics meet the requirements of the microcapsules. The preparation method and characterization of microcapsules can provide a theoretical basis and supporting data for their production and practical applications, and have a certain application value. At the same time, the use of 10% gibberellic acid microcapsules and 10% gibberellic acid soluble powder can reduce herbicide damage to crop growth caused by methyldisulfuron in wheat. The two types of microcapsules have a different effect. Through the indexes of microcapsule particle size, encapsulation rate, release rate, and the effect on wheat growth, it was found that the performance of the microcapsules prepared by the *in situ* polymerization method was better than that of the microcapsules prepared by the phase transfer method, and the microcapsules can prolong the product residual efficacy in the field *via* controlled release; therefore, the *in situ* polymerization method was recommended for practical production.

DATA AVAILABILITY STATEMENT

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

AUTHOR CONTRIBUTIONS

BZ was responsible for the framework design and topic selection of this whole manuscript. GZ designed and implemented solutions and wrote the article. JW and LL were responsible for experimental verification and data sorting. TM was responsible for data checking. YC is in charge of proofreading articles. All authors approved this manuscript for publication.

FUNDING

This work was funded by the Shanxi Department of Science and Technology, Shanxi Science and Technology Foundation Platform Project (201605D121024), and Shanxi Agricultural University Science and Technology Innovation Project (412579).

REFERENCES

Cao, B., Yang, Q., and Zhang, Y. (2018). Experimental study on the pesticide damage of herbicide methyldisulfuron-methyl in wheat field. *J. Agric. Sci. Technol.* 116–119. doi: 10.3969/j.issn.1007-5739.2018.11.070

Chen, F. L. (2015). *Processing and Application of New Pesticide Formulations*. Beijing: Chemical Industry Press. p. 26–34.

Chen, L. (2018). *Study on the Effect of Brassinolide on Rice Germination and Growth*. Changsha: Hunan Agricultural University.

Cui, J. B., Li, X., Mu, B., Guo, J. H., Cui, J. F., and Yang, B. P. (2018). Research progress on the application of microcapsules in polymer materials. *Chem. Mater.* 12, 1–4.

Jegat, C., and Taverdet, L. J. (2000). Stirring speed influence study on the microencapsulation process and on the drug release from microcapsules. *Polymer Bullet.* 44, 345–351. doi: 10.1007/s002890050612

Liu, P., Guo, Z. H., Chen, X., Guo, L. H., Tian, Y. B., Qiu, T. H., et al. (2015). Research progress on the role of gibberellin in wheat development and stress response. *Biotechnol. Adv.* 5, 253–258. doi: 10.3969/j.issn.2095-2341.2015.04.01

Liu, S. S., Zhang, Y. J., Xu, F., Chu, Z., and Na, Z. G. (2018). Optimization of preparation process and slow-release analysis of vanilla orchid essential oil microcapsules. *J. Trop. Crop Sci.* 39, 1423–1430. doi: 10.3969/j.issn.1000-2561.2018.07.024

Liu, X. G., and Xie, Y. T. (2005). Comparison of preparation methods of microcapsules. *Grain Food Indust.* 12, 28–30. doi: 10.3969/j.issn.1672-5026.2005.01.010

Wang, K. Y. (2009). *Pesticide Preparation Science*. Beijing: China Agricultural Press.

Wang, Y.u., Chen, T. B., Huang, C. Y., and Sun, B. H. (2000). 12% Huwei microcapsules for broadleaf weed control in soybean fields. *Pesticide.* 39, 32–33. doi: 10.3969/j.issn.1006-0413.2000.07.017

Xue, Y. (2019). *Study on the Preparation and Performance of a Midazolam Capsule Suspension*. Chongqing: Southwest University.

Ye, B., Xu, F. B., and Guo, J. W. (1996). *GB15955-1995 Gibberellin Drug.* Beijing.

Yu, Q. Q., Xia, M. L., Hu, W. H., Zhang, Y. Y., and Shi, T. S. (2019). Effects of several pesticide adjuvants on the penetration of gibberellin on wheat leaves. *Jiangsu Agri. Sci.* 47, 125–128. doi: 10.15889/j.issn.1002-1302.2019.22.028

Zhang, D. X. (2011). *Comparative Study of Urea-Formaldehyde Resin and Polyurethane Wall Material Microcapsule Pesticide Application Performance*, Taian: Shandong Agricultural University.

Zhang, G. S. (2012). *Application of Beta-Cypermethrin Microcapsule Suspension in Different Wall Materials and its Environmental Characteristics*. Taian: Shandong Agricultural University.

Zhang, G. S., Yang, F. H., Guo, H. L., Sang, J., Gao, Z. S., and Liu, F. (2012). Differences in migration and photolysis of beta-cypermethrin microcapsule suspensions in soils from different wall materials. *J. Pestic. Sci.*14, 214–220. doi: 10.3969/j.issn.1008-7303.2012.02.16

Zhao, D., Han, Z. R., Du, Y. C., Mu, W., and Liu, F. (2007). Microencapsulation and release characteristics of chlorpyrifos. *China Agri. Sci.* 40, 2753–2758. doi: 10.3321/j.issn:0578-1752.2007.12.013

Zhou, H. Z., Cui, Y., and Dong, G. X. (2019). Study on the stability of methyldisulfuron-dispersible oil suspending agent. *Pesticides.* 58, 799–801.

Conflict of Interest: YC was employed by Bayer Crop Science (China) Co., Ltd.

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

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: 13 July 2022
doi: 10.3389/fpls.2022.963170

Detecting Asymptomatic Infections of Rice Bacterial Leaf Blight Using Hyperspectral Imaging and 3-Dimensional Convolutional Neural Network With Spectral Dilated Convolution

Yifei Cao[1,2†], Peisen Yuan[1†], Huanliang Xu[1], José Fernán Martínez-Ortega[3], Jiarui Feng[1,2] and Zhaoyu Zhai[1]*

[1] College of Artificial Intelligence, Nanjing Agricultural University, Nanjing, China, [2] College of Engineering, Nanjing Agricultural University, Nanjing, China, [3] Departamento de Ingeniería Telemática y Electrónica (DTE), Escuela Técnica Superior de Ingeniería y Sistemas de Telecomunicación, Universidad Politécnica de Madrid (UPM), Madrid, Spain

OPEN ACCESS

Edited by:
Jian Su,
Nanjing University of Information Science and Technology, China

Reviewed by:
Zhixin Zhou,
Hangzhou Dianzi University, China
Xinyan Wang,
Xi'an Peihua University, China

***Correspondence:**
Zhaoyu Zhai
zhaoyu.zhai@njau.edu.cn

† These authors have contributed equally to this work and share first authorship

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

Received: 07 June 2022
Accepted: 22 June 2022
Published: 13 July 2022

Citation:
Cao Y, Yuan P, Xu H, Martínez-Ortega J, Feng J and Zhai Z (2022) Detecting Asymptomatic Infections of Rice Bacterial Leaf Blight Using Hyperspectral Imaging and 3-Dimensional Convolutional Neural Network With Spectral Dilated Convolution. Front. Plant Sci. 13:963170. doi: 10.3389/fpls.2022.963170

Rice is one of the most important food crops for human beings. Its total production ranks third in the grain crop output. Bacterial Leaf Blight (BLB), as one of the three major diseases of rice, occurs every year, posing a huge threat to rice production and safety. There is an asymptomatic period between the infection and the onset periods, and BLB will spread rapidly and widely under suitable conditions. Therefore, accurate detection of early asymptomatic BLB is very necessary. The purpose of this study was to test the feasibility of detecting early asymptomatic infection of the rice BLB disease based on hyperspectral imaging and Spectral Dilated Convolution 3-Dimensional Convolutional Neural Network (SDC-3DCNN). First, hyperspectral images were obtained from rice leaves infected with the BLB disease at the tillering stage. The spectrum was smoothed by the Savitzky–Golay (SG) method, and the wavelength between 450 and 950 nm was intercepted for analysis. Then Principal Component Analysis (PCA) and Random Forest (RF) were used to extract the feature information from the original spectra as inputs. The overall performance of the SDC-3DCNN model with different numbers of input features and different spectral dilated ratios was evaluated. Lastly, the saliency map visualization was used to explain the sensitivity of individual wavelengths. The results showed that the performance of the SDC-3DCNN model reached an accuracy of 95.4427% when the number of inputs is 50 characteristic wavelengths (extracted by RF) and the dilated ratio is set at 5. The saliency-sensitive wavelengths were identified in the range from 530 to 570 nm, which overlaps with the important wavelengths extracted by RF. According to our findings, combining hyperspectral imaging and deep learning can be a reliable approach for identifying early asymptomatic infection of the rice BLB disease, providing sufficient support for early warning and rice disease prevention.

Keywords: bacterial leaf blight, asymptomatic infection, hyperspectral imaging, deep learning, spectral dilated convolution, interpretable

INTRODUCTION

Rice is one of the most important grain crops, more than half of the world's population relies on it for food (Wu et al., 2020). Achieving steady and high rice yield has always been the goal of agricultural production. Bacterial leaf blight (BLB) disease, as one of the three major diseases of rice, is not evenly distributed in rice fields, but occurs in patches with large areas of the field free of disease in the early stages of infestation. In recent years, the outbreak area of BLB disease accounts for about one third of the total planting areas, and the average diseased plant rate is about 10%. It has a huge impact on the yield of rice. Generally, the yield is reduced by up to 50–60%, and even the grains are not harvested (Zarco-Tejada et al., 2018; Zhang et al., 2020). To reduce the negative effects of rice BLB disease, farmers used to treat a large number of pesticides. Overuse of pesticides not only increases the treatment expenditure, but also pollutes the environment (Šebela et al., 2018; Zhang J. et al., 2019). From the perspective of plant protection, the primary task is to quickly and accurately identify the potential occurrence of BLB, and then apply chemical treatments with the required amount (Tian et al., 2021). Therefore, the early identification of BLB is particularly critical. Early detection and prevention, as well as timely guidance, enable farmers to take efficient measures to control the spread of the disease, thereby reducing the amount of pesticides applied, and achieving the goal of sustainable agriculture (Rahman et al., 2020; Shu et al., 2021).

As computer vision and deep learning techniques have developed rapidly in recent years, they have shown great promise in detecting plant diseases (Qiu et al., 2021). Various rice disease detection methods have been proposed by detecting external changes of infected rice leaves from RGB images. Jiang et al. (2021) used a VGG-16 model and RGB images to recognize diseases of rice leaves and wheat leaves at the same time. Lu et al. (2017) proposed a deep convolutional neural network to identify a dataset of 500 RGB images containing 10 rice diseases, and the accuracy achieved at 95.48%. Besides, researchers have applied evolutionary approaches to neural architecture search of convolutional neural networks for improving computational efficiency (Xue et al., 2021a,b).

Although deep neural networks have achieved great success in detecting rice disease from RGB images, it is worth noting that these networks may fail to generate correct results for the early asymptomatic BLB disease detection based on RGB images. The examples of healthy and asymptomatic leaves are presented in **Figure 1**. Without giving labels in advance, one can hardly identify whether the leaf is healthy (**Figure 1A**) or asymptomatic (**Figure 1B**) due to their similar visual textures.

Since the rice BLB disease is caused by pathogen, rice leaves under such a stress would experience two periods, including asymptomatic and symptomatic stages (Deng et al., 2019; Tian et al., 2021). Although there are no significant lesions shown at the asymptomatic stage, the inner chemical composition has changed according to the plant pathology theory (Chen et al., 2020). Therefore, this inner change in rice leaves motivated us to adopt the hyperspectral imaging technique to detect the asymptomatic infection of the rice BLB disease and we obtained the following results shown in **Figure 2**. It can be seen that there are distinguishable features between the wavelengths from 378 to 1033 nm of hyperspectral images of healthy and infected rice leaves. This finding motivates us to use the hyperspectral imaging technique to detect the rice BLB disease at the asymptomatic stage, thereby providing an earlier warning to the farmers and assisting them in decision making about chemical treatments.

Hyperspectral imaging is a technique that analyzes multiple wavelengths of reflectance intensity of each pixel instead of just investigating primary colors (e.g., red, green, and blue) that only cover the wavelength range from 450 to 680 nm (Lowe et al., 2017; Zhang et al., 2021a). In comparison to traditional spectral and image technologies, hyperspectral imaging technology can obtain multiple wavelengths of spectrum and image information at the same time, and it has been demonstrated to be an effective and non-destructive technique for detecting crop diseases across multiple scales (Franceschini et al., 2019; Riefolo et al., 2021). Adopting the hyperspectral imaging technique to detect leaf disease has become a popular approach because hyperspectral imaging has high potential for finding new insights about plant diseases. Through the information fusion of multiple wavelengths, hyperspectral imaging can achieve better classification performance than using RGB images (Cabrera Ardila et al., 2020; Wang et al., 2021). Koushik et al. (2019) used hyperspectral imaging to detect soybean charcoal rot disease. Zhao et al. (2020) combined hyperspectral imaging and SVM to detect wheat early powdery mildew. Zhang et al. (2021a) applied in-situ hyperspectral imaging to diagnose the symptoms of sheath blight disease on rice stalk. Kaitlin et al. (2020) detected potato pre-symptomatic of late blight and early blight by hyperspectral imaging. From above works, it can be concluded that hyperspectral imaging is a powerful technique to detect the early disease of plants. Meanwhile, it is also worth noting that the high dimensionality of hyperspectral images might bring a huge challenge to the computational complexity, which is an urgent problem to be solved in the detection of plant diseases using hyperspectral images (Zhang et al., 2020).

Because of the high dimensionality of hyperspectral images, shallow machine learning models cannot perfectly handle the target detection tasks in hyperspectral images (Kaitlin et al., 2020). Deep learning, on the other hand, has shown its great potential in diverse applications. Hyperspectral images can be treated as a stack of 2D matrices, representing the correlations between spatial and spectral directions (Zhang J. et al., 2019). Therefore, many scholars have contributed to the combination of deep learning and hyperspectral imaging for plant disease identification (Polder et al., 2019; Xiao et al., 2022). Chu et al. (2022) proposed a shallow convolutional neural network with attention mechanism model to predict the early herbicide stress in wheat cultivars. Polder et al. (2019) designed a new imaging setup consisting of a hyperspectral line-scan camera and applied a convolutional neural network for detecting potato virus Y. Nguyen et al. (2021) applied both deep learning and machine learning models to identify grapevine early vein-clearing virus in hyperspectral images. The result showed that the former achieved a better classification result. Conclusively, combining

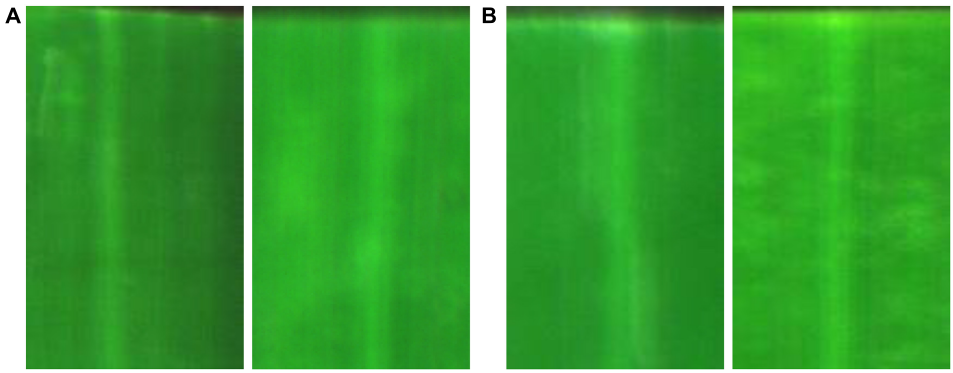


FIGURE 1 | RGB images of rice leaves: **(A)** healthy; **(B)** asymptomatic.

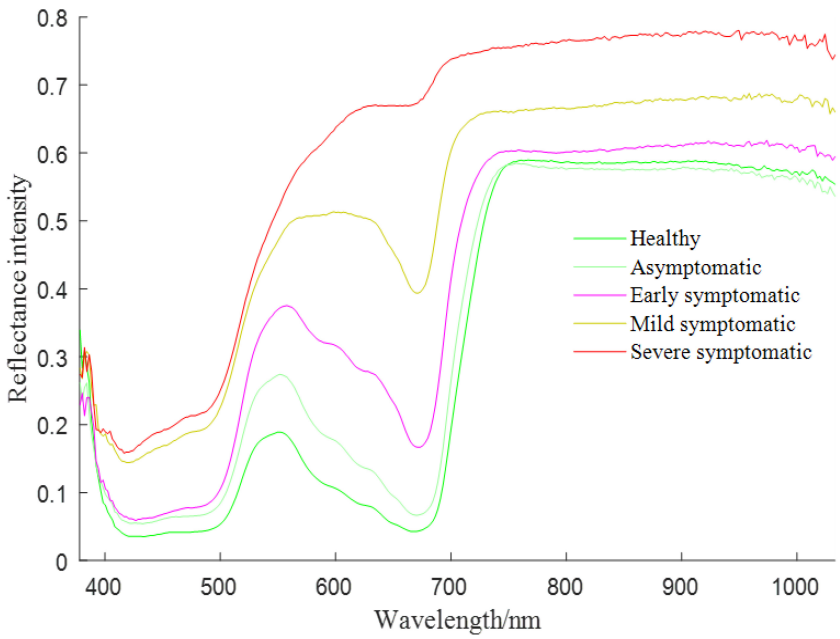


FIGURE 2 | Spectra of infected leaves under different stages.

deep learning with hyperspectral imaging is very promising in the plant disease detection task.

The purpose of this paper is to apply hyperspectral imaging and deep learning to detect asymptomatic infections of the rice BLB disease. The main contributions of this paper were summarized as follows: (1) to explore the applicability of using random forest (RF) and principal component analysis (PCA) to extract sensitive features from raw hyperspectral data, thereby improving the computational efficiency; (2) to build a spectral dilated convolution 3-dimensional convolutional neural network (SDC-3DCNN) model and test the effect of the number of input features on the detection performance; (3) to assess the detection performance of the SDC-3DCNN model under different spectral dilated ratios (SDR); (4) to interpret the important wavelengths with a saliency map method. In conclusion, the proposed SDC-3DCNN model is able to detect the asymptomatic infection of the rice BLB disease, thereby providing early warnings before the disease outbreak. This result may assist in arranging chemical treatments for disease control.

MATERIALS AND METHODS

Experimental Materials

Rice Planting and Artificial Inoculation

The experimental materials were processed and collected in 2020 at a greenhouse base in Nanjing Agricultural University, Nanjing, Jiangsu, China. The rice seeds were sown and then transplanted into plastic pots (three plants per pot) on June 17th, 2020. The rice was grown in plastic pots (35 cm diameter × 32 cm height) filled with paddy clay soil. A total of 50 pots were used for the hyperspectral data collection. Data collected from

different rice leaves provided an opportunity to evaluate the reproducibility and reliability of the findings on disease detection. To ensure consistent management practices in the greenhouse environment, the basal nutrition fertilizers (nitrogen provided by urea, 150 kg·ha^{-2}; P_2O_5, 135 kg·ha^{-2}; and K_2O, 18.3 kg·ha^{-2}) were applied prior to transplantation, and a second nitrogen topdressing (N, 150 kg·ha^{-2}) was applied during the tillering stage. The rice plants were irrigated as needed to ensure that the soil in each pot was always covered by a shallow layer of water. All plants were placed outdoors and were not transferred to the greenhouse until 1 week before the inoculation treatment. The greenhouse comprised two layers of transparent materials and was equipped with air conditioning and humidifying facilities to provide suitable environmental conditions (26–32°C, over 90% relative humidity, and a photoperiod of 14 h) for the artificial inoculation of the BLB pathogen. The artificial inoculation operations were conducted on rice leaves to induce BLB infection at the tillering stage. The BLB pathogen (*Xanthomonas oryzae pv.oryzae Xoo*) was provided by the College of Plant Protection at Nanjing Agricultural University, Nanjing, Jiangsu, China. After the BLB pathogen was isolated, it was transferred to the plate medium (beef extract 0.3%, meat peptone 1%, sucrose 1%, and agar 2%). The bacteria were placed in an incubator at 28°C and cultivated for 48 h, diluted with phosphate buffered saline. The concentration was diluted to about 9×10^9 bacteria per milliliter by the turbidimetric method. The surgical scissors were used to dip the bacterial solution and cut off the tip of the rice flag leaf to complete the bacterial inoculation. All inoculation operations were completed within 3 h. After inoculation, all plants were completely covered with black, light-tight plastic materials for 48 h. Temperature (26–32°C), relative humidity (≥90%), and light conditions were all strictly controlled to ensure a successful infection.

Definition of the Disease Infection Process

Previous studies have typically defined disease severity at the leaf level as the average percentage of symptomatic surface areas (Bock et al., 2020; Tian et al., 2021). However, rice leaves with asymptomatic infections cannot be described by this method. For the rice BLB disease, the appearance of obvious symptoms on rice leaves indicates that a large area has begun to spread, which is not conducive to the timely prevention of the disease. Effectively identifying asymptomatic infections is crucial for the prevention and control of the BLB disease.

Since the spectrum of health pixels is different from infected areas, a pixel-level annotation method can be used to visualize areas of asymptomatic infection and to more precisely define the disease levels (Lowe et al., 2017). Labels for each pixel sample were determined by combining visual inspection of leaf color with spectral changes (Sun et al., 2018). Specifically, the healthy ROI was defined as the same uninoculated area without any change in the spectrum after inoculation (**Figure 3A**). The asymptomatic ROI was defined as the area where obvious disease lesions had not yet appeared but the spectral of infected leaves changed. In the asymptomatic stage, light yellowish-green watery lesions could be observed on a few infected leaves, but they are often not easily noticed under field conditions (**Figure 3B**).

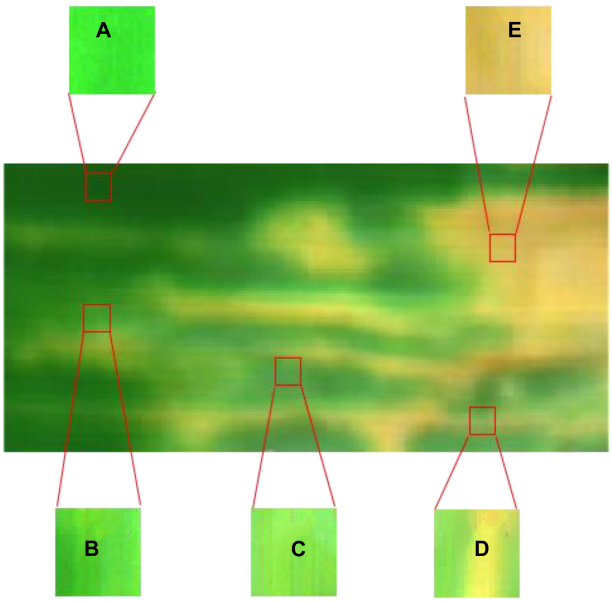


FIGURE 3 | The pixel-level annotation of disease severity levels. **(A)** Denotes pixels that belong to the healthy class. **(B)** Denotes pixels that belong to the asymptomatic class. **(C–E)** Denote pixels that belong to the symptomatic class.

The symptomatic ROI was defined as the leaves are turning into wavy yellow green or gray green spots along one or both sides of the leaf margin or along the mid-vein, along with irregular chlorotic spots, and then turning into withered yellow patches or large patches (**Figures 3C–E**). The boundary between the symptomatic part and the asymptomatic part is observable, whereas it is difficult to distinguish between the asymptomatic and healthy pixels.

Data Collection and Processing

The configuration of the hyperspectral imaging system and parameter settings in this study can be referred to Zhang et al. (2015). The correlative settings of HSI, including the speed of the motor and positions, could be set by the software (Isuzu Optics Corp, Taiwan, China). The main components and parameters of the hyperspectral imaging system are shown in **Table 1**.

The imaging system was pre-heated for 5 min before collecting hyperspectral images. For data collection, rice leaves were placed on the stage below the imaging lens at an object distance of 27 cm, and leaves were fixed on a black background (**Figure 4**).

The spectrum was recorded in the wavelength range of 378.28–1033.05 nm by the hyperspectral imaging system. In the first conversion step, the information was transformed to a cube format containing the spatial information in the x–y directions and the spectral information in the z direction (Bauriegel et al., 2011). Due to the existence of dark current in the CCD camera and the uneven intensity distribution of the light source in different spectra, the obtained hyperspectral image is unstable. Therefore, the hyperspectral image correction was carried out by the black and white correction method (Zhang et al., 2015). After

TABLE 1 | Main components of the imaging system and parameter settings.

Parameters	Value (unit)	Parameters	Value (unit)
Spectral camera	Raptor EM285	Light input	21V/200W halogen light
Dispersion	97.5 nm/mm	Spectral range	378.28–1033.05 nm
Spectral resolution	2.14 nm	Image size	1632 × 1401 pixels
Spatial resolution	Spot diameter < 9 μm	Object distance	27 cm
Aberration	Halo < 1.5 μm, trapezoid < 1 μm	Exposure time	8 ms
Aperture	F/2.4	Move speed	0.8 mm/s

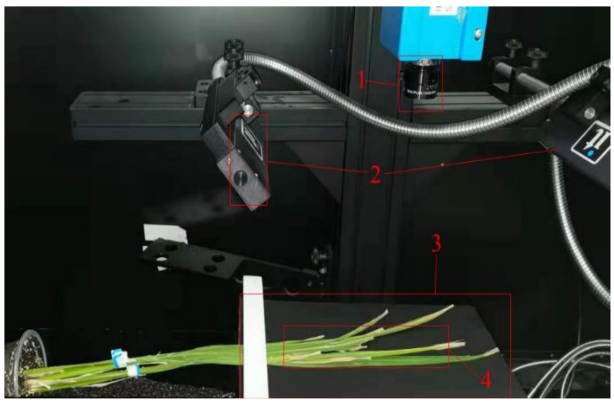


FIGURE 4 | The schematic of the hyperspectral imaging system. 1. Imaging lens; 2. halogen lamps; 3. transportation platform; 4. leaves sample.

collecting the hyperspectral images of all diseased rice leaves, the leaf samples were quickly put back into the original environment for further cultivation so as to facilitate the next image collection while reducing the impact of the environment on plant growth and disease development.

Overview of Data Processing and Modeling Pipeline

The spectrum used in this study ranged from 378.28 to 1033.05 nm, 306 wavelengths in total. In order to reduce redundancy and increase the computational efficiency, RF and PCA were used to extract spectral features from raw hyperspectral images. According to the number of extracted features, nine datasets were established. The detection performance of the SDC-3DCNN model over the nine datasets was tested. Meanwhile, the effect of dilation ratios was also examined through experiments. Lastly, the saliency score of wavelength channels was calculated and sorted for interpretation of feature importance. The overview workflow of the data processing and modeling pipeline for early rice BLB disease detection is shown in **Figure 5**.

Hyperspectral Preprocessing

Raw hyperspectral images were analyzed with the software ENVI 5.3 (ITT Visual Information Solutions, Boulder, United States). With this software, symptomatic and healthy pixels of leaves could easily be labeled in a false color image at 450, 550, and 650 nm by visual inspection. While asymptomatic pixels have to be labeled according to the spectral change. The false color images also facilitated the proper manual setting of ROIs and the selection of tissues for spectral analysis. In all hyperspectral images, spectra of healthy and infected tissue areas were obtained pixel-wise. To reduce the impact of noise at both ends of the spectrum, only 450–950 nm (a total of 232 wavelengths) were intercepted for further analysis. The original spectra of different ROIs are shown in **Figure 6A**.

In order to eliminate the random noise in the spectral signal and improve the signal-to-noise ratio of the sample signal, the Savitzky–Golay (SG) smoothing filter was used to reduce the random noise (the kernel size was 5 × 5 × 5, the polynomial order was 3, and the filter value was calculated at the central node of the kernel). The noise in the spectra of different ROIs was greatly reduced after SG smoothing, and the interference of size difference and different information structures was removed (**Figure 6B**).

The original spectra and SG smoothed spectra in the range of 450–950 nm are shown in **Figure 6**. The general distribution of the reflection intensity of the healthy ROI is consistent with the asymptomatic ROI, however, both of which are different from the reflection intensity of the symptomatic ROI. From 450 to 550 nm, the reflectance of all spectra shows an increasing trend. While the trend changes between 550 and 750 nm, for healthy and asymptomatic ROI, the reflectance shows a significant decrease in the spectral range of 550–680 nm and a rapid increase in the range of 680–750 nm. The reflectance of the asymptomatic spectrum is higher at each wavelength than the reflectance of the healthy spectrum. For symptomatic ROI, the reflectance from 680 to 750 nm increases less than the reflectance of asymptomatic ROI. Finally, between 750 and 950 nm, the spectrum tends to be flat. The difference between different spectra provides a basis for classifying different pixels based on the spectral information.

Features Extraction by Different Algorithms

Considering that the variation of disease rice physiological parameters could induce strong responses at specific spectral wavelengths, the unbalance of the spectral features selected from the 450–950 nm range could be attributed to the sensitivity of different physiological parameters to disease infection. Spectral information covers wavelengths from 450 to 950 nm and is characterized by a high dimension of redundancy between adjacent wavelengths. Excessive redundant spectral information brings great challenges to detection methods and computational complexity (Hennessy et al., 2020; Zhang et al., 2021b). Therefore, it is necessary to compress the amount of data by a dimensionality reduction method to reduce the cost of subsequent processing on the basis of not dropping the effective feature spectral information (Poona et al., 2016; Sadeghi-Tehran et al., 2021).

On the one hand, the principal component analysis (PCA) algorithm is a common data compression method and it is often used for dimensionality reduction of high-dimensional data. PCA can extract principal components (PCs) of original hyperspectral data (Hsieh and Kiang, 2020). The main idea of

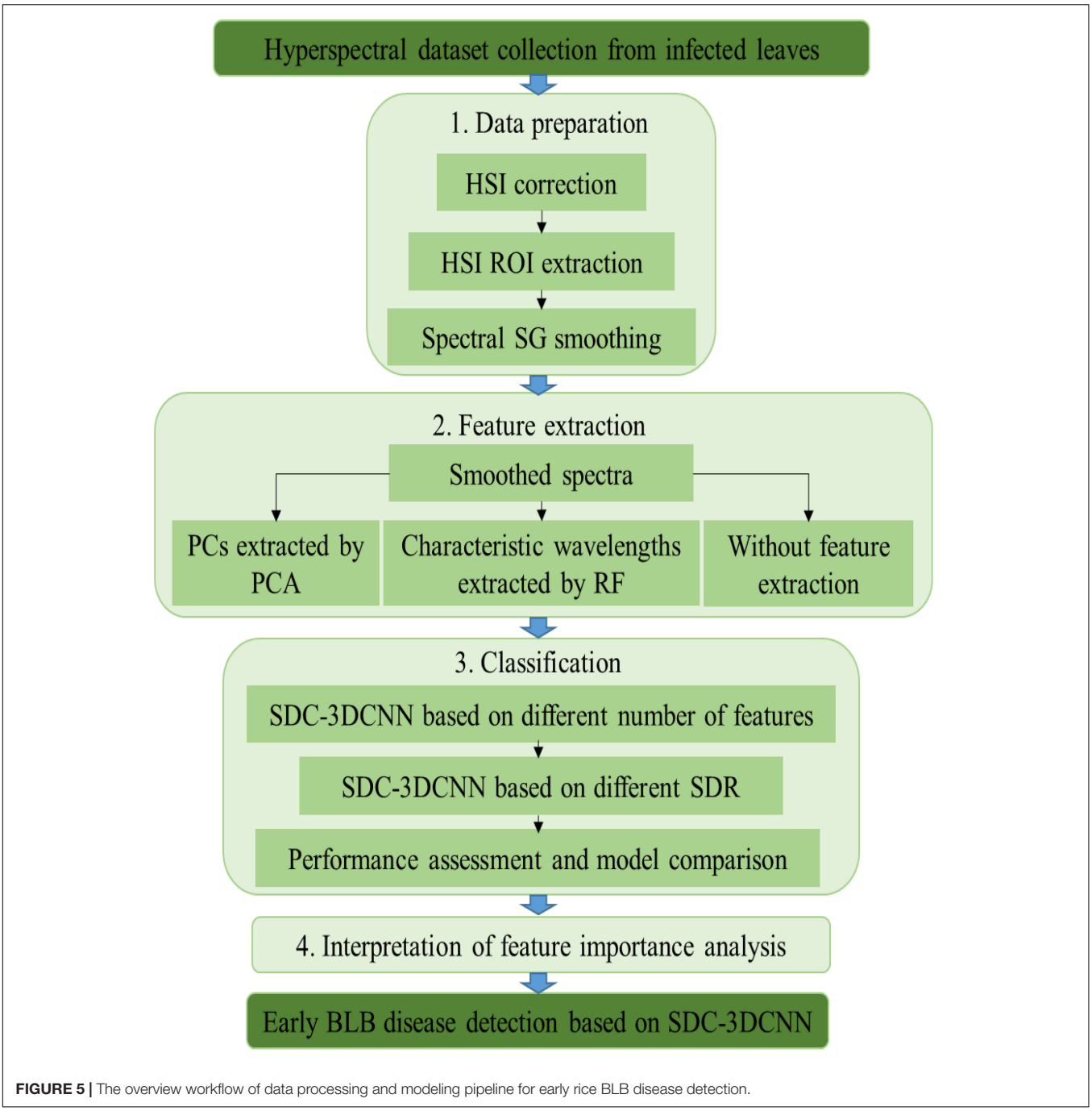


FIGURE 5 | The overview workflow of data processing and modeling pipeline for early rice BLB disease detection.

PCA in this paper is to map 232 wavelengths to a k-dimensional feature space (k < 232). On the contrary to simply removing less important dimensional features from the original space, these k-dimensional features can be viewed as PCs through mapping the features in the original space into a latent space. The PCs were obtained by following the calculation process shown in **Figure 7**.

The PCA was used to extract PCs from the smoothed 232 wavelengths of hyperspectral images. The original hyperspectral data was projected to a k-dimensional space for dimensionality reduction.

On the other hand, the random forest wavelength extraction method is to randomly replace each wavelength and to evaluate its importance (Ma et al., 2018). The higher the wavelength importance, the greater the variation in the prediction error rate would achieve (Hidayat et al., 2017; Speiser et al., 2019). In this way, each wavelength was scored by the change in the error rate of the out-of-bag data before and after wavelength replacement, so as to obtain the characteristic wavelength with a high importance score (Tan et al., 2020). In this paper, in order to reduce the dimension of original hyperspectral data, a different

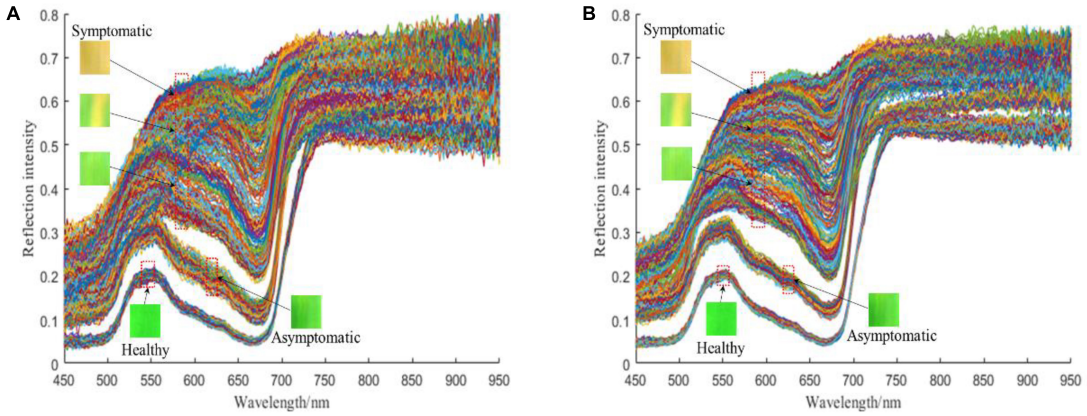


FIGURE 6 | The reflection intensity of different ROIs: **(A)** Original spectra; **(B)** SG smoothed spectra. A single line represents the reflection intensity of an individual spectrum. The total number of spectra in **(A)** and **(B)** is both 1000.

number of characteristic wavelengths were selected according to the important score.

The random forest algorithm used for extracting characteristic wavelengths is described as follows. The number of decision trees in the RF is denoted by N_{tree}, set at 200; the number of the original wavelength is denoted by d, set at 232. The wavelength importance measurement based on out-of-bag error analysis of a single feature $X^j(j = 1, 2, , d)$ was calculated as follows:

Step1: Calculate the number of Out-of-Bag error samples $ErrOOB_i$ of Out-of-Bag data OOB_i corresponding to the i^{th} decision tree.

Step2: Keep other wavelengths unchanged, randomly change the wavelength X^j in OOB_i, and obtain OOB_i^j.

Step3: Re-calculate the number of out-of-bag error samples $ErrOOB_i^j$ of out-of-bag data OOB_i^j.

Step4: Repeat steps 1, 2, and 3 to obtain $\left\{ErrOOB_i^j | i = 1, 2, , N_{\text{tree}}\right\}$.

Step5: The importance score $VI\left(X^j\right)$ of wavelength X^j was calculated by Equation 1:

$$VI(X^j) = \frac{1}{N_{tree}} \sum_{i}^{N_{\text{tree}}} (ErrOOB_i^j - ErrOOB_i) \quad (1)$$

Both PCA and RF were adopted to reduce the dimension of raw hyperspectral data, thereby minimizing the computational complexity. The importance score of each wavelength from the original spectrum was ranked by the above methods, and the top 50, 100, 150, and 200 wavelengths extracted by RF were used to establish four datasets. On the other hand, the principal components (PCs) of the top 50, 100, 150, and 200 rankings extracted from the original spectrum by PCA were used to establish another four datasets. In addition, the original smoothed 232 wavelengths without feature extraction were used to establish the ninth dataset for comparison as well. In total, nine datasets were established.

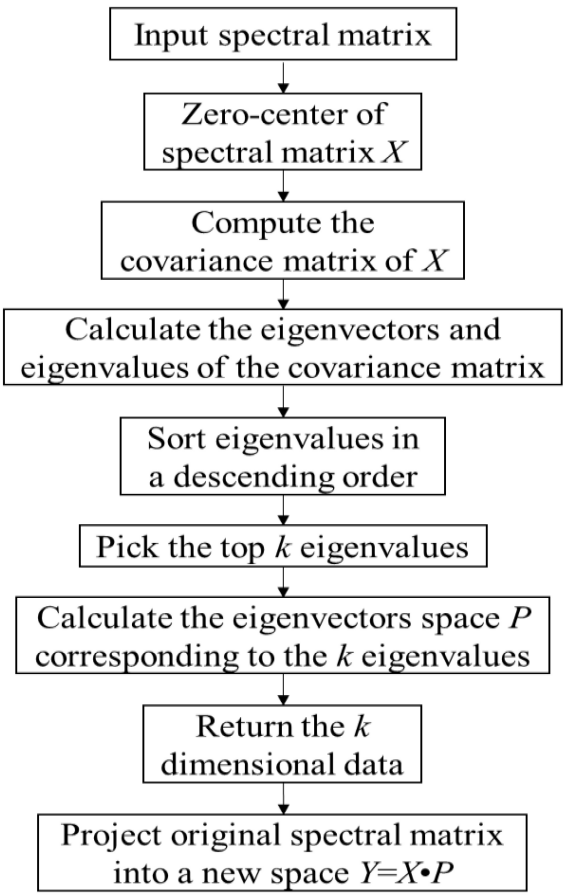


FIGURE 7 | Flowchart of extracting the PCs.

Spectral Dilated Convolution

Dilated convolution is usually applied to expand the receptive field without changing the original structure or the number of parameters of the model (Cao and Guo, 2020). Whereas traditional 2D dilated convolutions can only enlarge the receptive

field along spatial dimensions (Xu et al., 2021). Due to the 3D character of hyperspectral data, 3-Dimensional Spectral Dilated Convolution (3D SDC) was developed to expand the receptive field along three dimensions of hyperspectral (**Figure 8**).

The receptive field of 3D SDC is decided by the spectral dilated ratio (SDR) as:

$$R_f = 2 \times (r_{SDR} - 1) \times (k - 1) + k \quad (2)$$

Where R_f represents the receptive field of a single convolution kernel; r_{SDR} represents the spectral dilation ratio; k represents the size of convolution kernels. Here k was set to 3 (Mohanty et al., 2016; Albattah et al., 2022).

As shown in **Figure 8**, the black cubes represent the convolution kernel, and the white cubes cover the receptive field. When SDR is set to 1, 2, and 3, the receptive fields of the convolution kernels are $3 \times 3 \times 3$, $7 \times 7 \times 7$, and $11 \times 11 \times 11$, respectively.

Residual Module

The increase in the model depth usually improves the performance of the neural network (Xu et al., 2021). However, such an increase may cause gradient vanishing or gradient explosion (Zhong et al., 2018). Cao and Guo (2020) suggested that the 3D residual connection can solve this problem.

The input information can be directly passed to subsequent layers through the residual module. The shortcut connections can be seen as identity mapping. In a residual block, the output of the l^{th} block is computed by Equation 3:

$$x_{l+1} = F(x_l) + h(x_l) \quad (3)$$

Where x_l and x_{l1} are the input and output of the l^{th} block, respectively. $F(x_l)$ is a residual mapping function, and $h(x_l)$ is an identity mapping function.

As shown in **Figure 9**, the shortcut connection in the l^{th} residual block, shown in blue background, $h(x_l)$ is basically a direct connection between the input and output of the l^{th} block, while $F(x_l)$ usually contains multiple convolution layers and batch normalization and activation.

Spectral Dilated Convolution 3-Dimensional Convolutional Neural Network Framework for Early Disease Detection

In this paper, we compared different numbers of input features. The RF can sort the original wavelength according to the importance score. The top 50, 100, 150, and 200 characteristic wavelengths extracted by the RF were treated as inputs to the SDC-3DCNN model, respectively. On the other hand, the top 50, 100, 150, and 200 PCs extracted by PCA from the original spectrum were also treated as inputs to the SDC-3DCNN model, respectively. In addition, the smoothed 232 wavelengths were treated as the input to the SDC-3DCNN model for comparison. The framework of the SDC-3DCNN model is presented in **Figure 10**, where H, W, and D represent the height, width, and size along the spectral dimension of the data cube. The SDC modules can extract and fuse features corresponding to multiple wavelength resolutions, so the important wavelength information can be more effectively used. We also employed residual blocks to avoid the gradient vanishing problem (Xu et al., 2021). The target sample from the infected rice leaf HSI is composed of 31680 (132×240) pixels, which are divided into three classes: healthy, asymptomatic, and symptomatic. The datasets were divided into a training set, a verification set, and a test set under the ratio of 8:1:1.

The 3D convolution is achieved by using a 3D kernel to stack multiple contiguous wavelengths into a cube (Pi et al., 2021). By this construction, the feature maps in the convolution layer are connected to multiple wavelengths frames in the previous layer, thereby capturing spectral information. Formally, the value at position (x, y, z) on the j^{th} feature map in the i^{th} layer, denoted as $v_{\text{ij}}^{\text{xyz}}$, is computed by Equation 4.

$$v_{ij}^{xyz} = f\left(\sum_{\text{m}} \sum_{h=0}^{h_i-1} \sum_{w=0}^{w_i-1} \sum_{d=0}^{d_i-1} w_{ijm}^{hwd} v_{(i-1)}^{(x+h)} m^{(y+w)(Z+d)} + b_{ij}\right) \quad (4)$$

Where H_i, W_i and D_i are the height, width, and the size along the spectral dimension of the kernel. $w_{\text{ijm}}^{\text{xyz}}$ is the value at the position (h, w, d) of the kernel connect to the m^{th} feature map in the

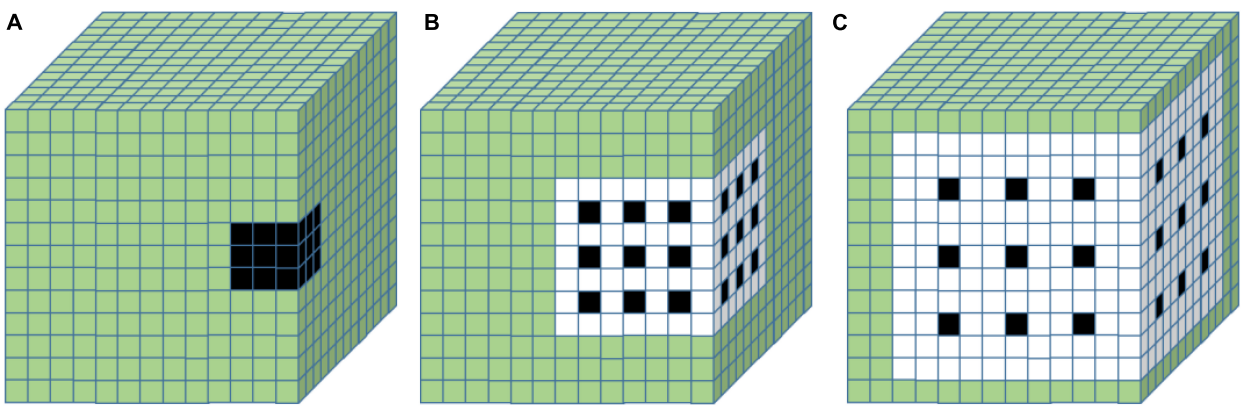


FIGURE 8 | 3D dilated convolution under different SDRs: **(A)** SDR = 1; **(B)** SDR = 2; **(C)** SDR = 3. It can be considered that zeros are inserted into the kernel to evenly space the filter under different SDRs.

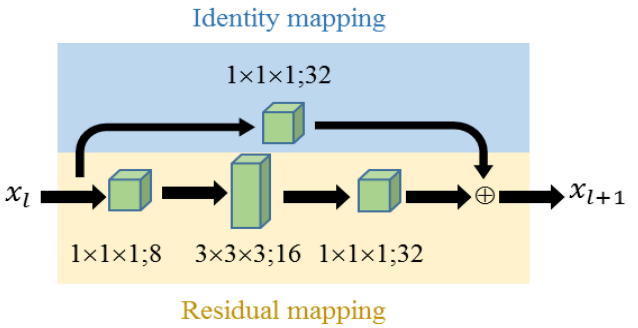


FIGURE 9 | Schematic of the 3D residual module.

previous layer. b_{ij} is the bias for this feature map. m indexes over the set of feature maps in the $(i-1)^{th}$ layer connected to the current feature map, and the other parameters are the same as in the 2D convolution (Ji et al., 2013).

The SDC-3DCNN model used cross entropy as the loss function and the stochastic gradient descent optimizer for training. The specific parameters of SDC-3DCNN were as follows: learning rate was set to 1×10^{-3}, weight decay coefficient was set to 1×10^{-6}, momentum was set to 0.95, epsilon was set to 1×10^{-5}, epoch was set to 50, dropout was set to 0.45. In order to achieve fast convergence, the training set was divided into multiple batches, and the batch size was set to 64.

Saliency Wavelengths in Hyperspectral Images

In order to prove the reliability of the detection results, we added post-hoc explanations. A saliency map method was applied to look into the classification results and further improve the overall design of the system. Saliency scores have been used as a popular visualization technique to detect how and why a deep learning neural network makes certain predictions (Zhang et al., 2021c).

In this research, a saliency explanation of wavelength channels was applied. Through one single back-propagation, the derivative ω from a specific predicted result to the input wavelength can be obtained from the well-trained model with the index $h(i, j, c)$. Here, (i, j) indicates the spatial arrangement of elements in ω, while c indicates spectral channels (Zhang et al., 2021c). The wavelength saliency scores W_c can be calculated by Equation 5.

$$W_c = \sum_i \sum_j |\omega_{h(i,j,c)}| \tag{5}$$

All contributions from one specific wavelength channel were summed. The saliency score represents the contribution of different wavelength channels to the classification result. The importance of a specific wavelength can be quantified by the saliency gradient magnitude at that wavelength (Nagasubramanian et al., 2018).

Classification Assessment

This paper considers the widely acknowledged criteria to evaluate the proposed SDC-3DCNN model, including precision, recall, F1 score, accuracy, and kappa coefficient. True positive, false positive, false negative, and true negative are denoted by TP, FP, FN, and TN, respectively. The formulas of precision (P) and recall (R) are presented in Equations 6 and 7.

$$\mathrm{P} = \frac{\mathrm{TP}}{\mathrm{TP} + \mathrm{FP}} \times 100\% \tag{6}$$

$$\mathrm{R} = \frac{\mathrm{TP}}{\mathrm{TP} + \mathrm{FN}} \times 100\% \tag{7}$$

It is expected that the values of precision and recall should be higher, but they are incompatible. The F1 score is a better metric that combines the characteristics of precision and recall to evaluate the model for different classes in the dataset. A high F1 score is also indicative of satisfactory classification performance.

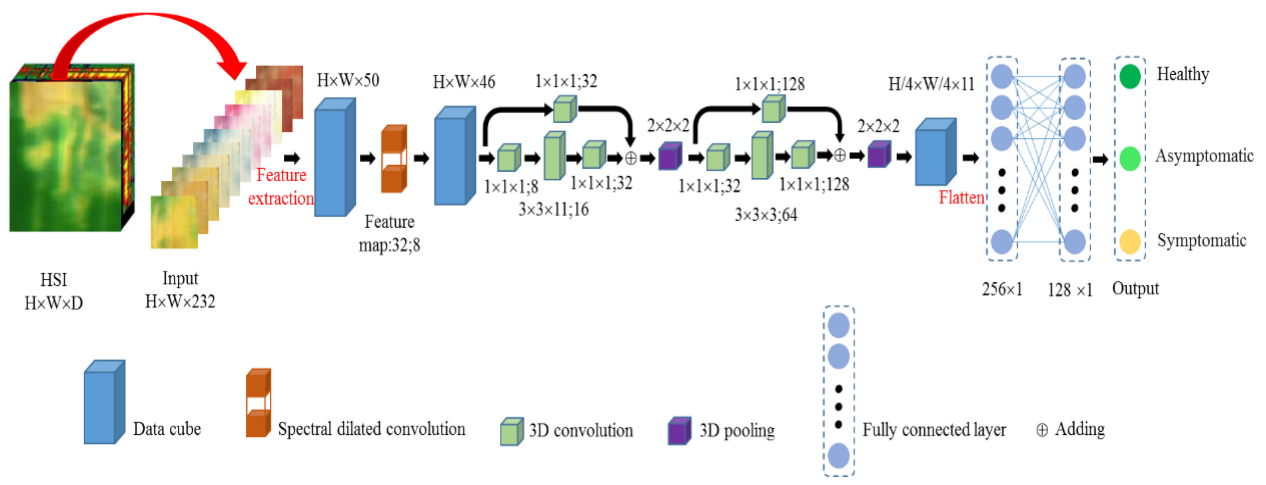


FIGURE 10 | Framework of the spectral dilated convolution 3-dimensional convolutional neural network model.

The F1 score formula is presented in Equation 8.

$$F1 = \frac{2 \times TP}{2 \times TP + FN + FP} \times 100\% \quad (8)$$

Accuracy (A) is another evaluation metric. In general, the higher the accuracy, the better effect that a model would achieve. The formula of accuracy is presented in Equation 9.

$$A = \frac{TP + TN}{TP + TN + FN + FP} \times 100\% \quad (9)$$

Kappa coefficient (K) is used for consistency test and classification accuracy. The higher the Kappa coefficient, the better consistency that a model would achieve. The formula of Kappa coefficient is presented in Equations 10 and 11:

$$K = \frac{A - P_e}{1 - P_e} \times 100\% \quad (10)$$

$$P_e = \frac{\sum_{i=1}^{n} a_{is} \times a_{si}}{N^2} \quad (11)$$

Where a_{is} is the sum of the elements in the i row of confusion matrix, a_{si} is the sum of the elements in the i column of confusion matrix, n is the number of columns in confusion matrix, N is the total number of pixel samples.

Besides, the training time and the number of trainable parameters of the model are also important indicators to evaluate the complexity of the model.

RESULTS

The proposed model is programmed in Python and implemented based on the Tensorflow and Keras open-source deep learning framework. The operating platform hardware configuration includes the NVIDIA GeForce RTX 2080Ti GPU and the AMD Ryzen 5-1600 Six-Core processor @ 3.20 GHZ CPUs.

A total of nine datasets were constructed from 50, 100, 150, and 200 characteristic wavelengths extracted by RF (4 datasets), 50, 100, 150, and 200 PCs extracted by PCA (4 datasets), and smoothed 232 wavelengths without any extraction (1 dataset). Each dataset is divided by 80% for training, 10% for validation, and 10% for testing. The performance of the SDC-3DCNN was evaluated as follows. First, the SDC-3DCNN model was trained and evaluated with those nine datasets. It is noted that at this stage, the SDR was set to 1, meaning that the SDC module did not have any effect. Second, we further explored the effect of the SDC module by adjusting the spectral dilated ratios. Finally, the saliency map method was used to interpret the important features that contributed the most to the classification results.

Features Extracted by Random Forest and Principal Component Analysis

In this paper, the RF was used to rank the importance of the original 232 wavelengths of the hyperspectral spectrum (**Figure 11A**). It can be seen that wavelengths with the highest importance scores are mainly distributed at 530–710 nm. It has been shown clearly in **Figure 11B** that the top 10 wavelengths with high importance scores are 547.2, 534.5, 551.4, 566.2, 697.4, 530.3, 693.0, 543.0, 538.7, and 568.4. The original 232 wavelengths were sorted according to the importance score obtained from the RF. The top 50, 100, 150, and 200 characteristic wavelengths were used to construct datasets for the detection of the asymptomatic BLB disease.

On the other hand, the PCA was used to compute the variance contribution of PCs. It can be seen from **Figure 12** that the x-axis represents the first principal component score and the y-axis represents the second principal component score. It can be seen that healthy, asymptomatic, and symptomatic pixels were projected into different categories. The top 50, 100, 150, and 200 PCs extracted by PCA were used to establish the datasets, respectively.

Detection Performance of Spectral Dilated Convolution 3-Dimensional Convolutional Neural Network With Different Input Features

Table 2 shows the performance comparison of using nine datasets to detect healthy, asymptomatic, and symptomatic pixels. The original 232 wavelengths without feature extraction, the top 50, 100, 150, and 200 PCs extracted by PCA, and the top 50, 100, 150, and 200 characteristic wavelengths extracted by RF were used as inputs to the SDC-3DCNN model, respectively. It can be seen from **Table 2** that the fewer the input features, the shorter the training time and the less the trainable parameters will be. For the same number of features extracted by RF and PCA, there was a minor difference between the training times. When the number of input features was reduced to 50, the training time is almost reduced to a quarter, and the number of trainable parameters is reduced to about 1/6, compared with the full-feature input. In terms of classification accuracy, when the characteristic wavelengths extracted by RF were used as the input, the overall performance is higher than that when the PCs extracted by PCA and the original 232 wavelengths without features extraction are used as the input. When the input is 50 characteristic wavelengths extracted by RF, the performance of the model achieved the best, with an accuracy of 94.7640% and a kappa coefficient of 92.1466%.

Figure 13 shows the accuracy (**Figure 13A**) and loss value (**Figure 13B**) of the validation set when inputting different features extracted by PCA and RF, as well as the full wavelengths. It can be seen that the worst accuracy was achieved when the full wavelengths were used as the input to the SDC-3DCNN model. However, using features extracted by PCA and RF as inputs, the SDC-3DCNN model achieved better results, especially the latter one.

From the results in **Table 2** and **Figure 13**, it can be seen that the performance is better when the characteristic wavelengths extracted by RF are treated as the input. This shows that the redundant spectral information not only increases the computational complexity and training time of the model but also reduces the performance of the detection model. It can be concluded that RF is more suitable for dimensionality reduction.

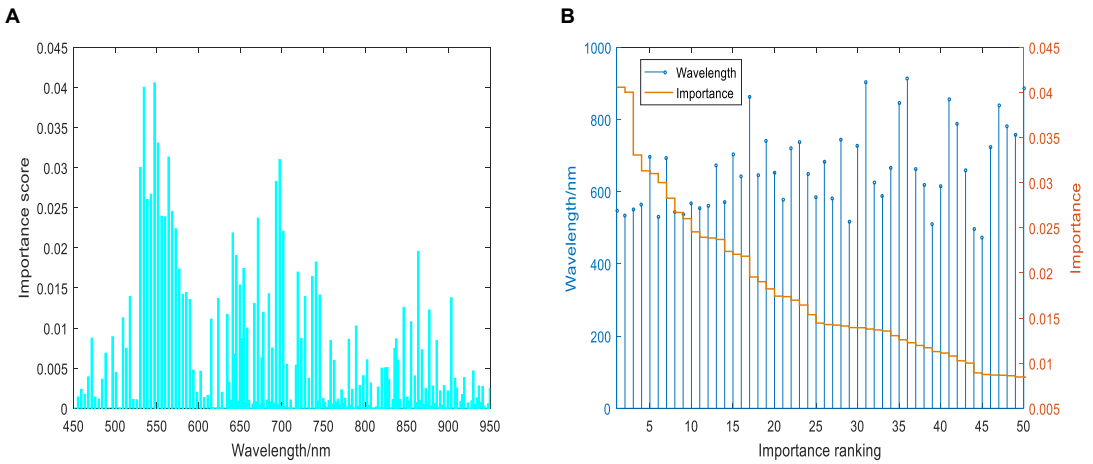


FIGURE 11 | The importance scores of wavelengths: **(A)** importance scores of 232 wavelengths; **(B)** the ranking of importance scores of the top 50 wavelengths.

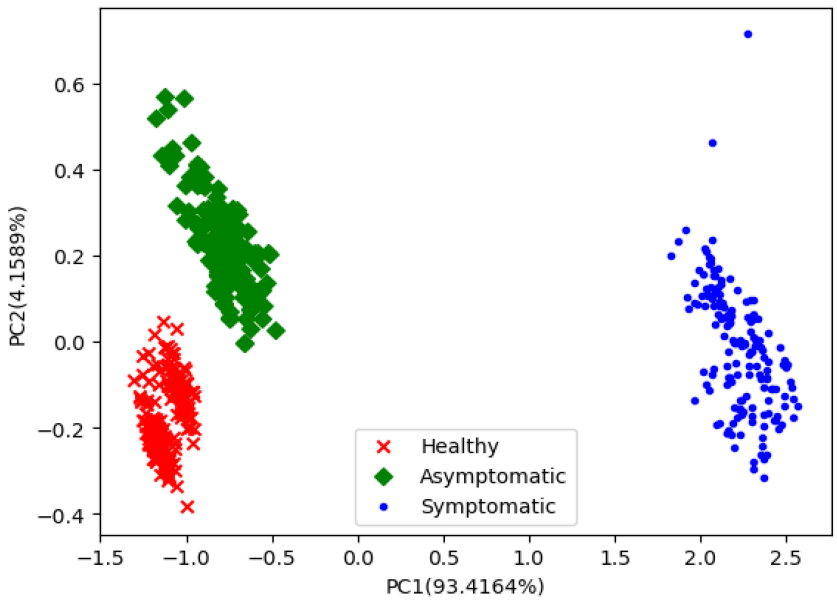


FIGURE 12 | PCA scores map of first two principal components.

Classification Performance Assessment With Different Spectral Dilated Ratios

It can be seen from **Table 2** that when the number of extracted features is 50, the model performance is the best. As a result, we set the number of input features to 50 and then tested the SDR module with various spectral dilated ratios. As shown in **Table 3**, for the 50 characteristic wavelengths extracted by RF, when the SDR is 5, the SDC-3DCNN model reached the optimal detection performance with an accuracy of 95.4427%. It is also worth noting that when the SDR is set to 5, the SDC-3DCNN model could detect healthy samples more accurately, with a precision of 97.7513%. The performance of detecting symptomatic samples ranks the second, with a precision of 94.3126%. It is nature that the SDC-3DCNN model achieved the worst precision (94.2804%) in the asymptomatic detection task because the asymptomatic samples just had minor differences with the healthy ones. Lastly, it is also glad to notice that the detection performance is generally improved when the SDR is enabled (when SDR is set greater than 1).

It can be seen from **Table 4**, for 50 PCs extracted by PCA as the input, when the SDR is 3, the detection performance reached the best, and the accuracy achieved at 93.2252%. Diving into the detection performance of each class, it is noted that the SDC-3DCNN model still can predict the healthy samples with the highest precision (96.6517%), following the symptomatic class with a precision of 93.2132%. This finding indicated that the class

TABLE 2 | Classification performance of different inputs.

Feature extraction method	Number of input features	Training time/s	Number of trainable parameters	Accuracy/%	Kappa coefficient/%
RF	**50**	**154.1598**	**1,559,731**	**94.7640**	92.1466
	100	263.6340	3,263,667	94.6772	92.0167
	150	382.1071	4,836,531	94.3261	91.4893
	200	502.7476	6,540,467	94.7122	91.0660
PCA	50	154.7476	1,466,515	92.9688	89.4517
	100	262.8174	2,870,451	92.6847	89.0279
	150	383.3145	4,443,315	92.6491	88.9721
	200	501.7399	6,147,251	92.9648	89.4443
None	232	573.6130	7,589,043	89.4492	84.1938

The bold values denote the achieved best performance.

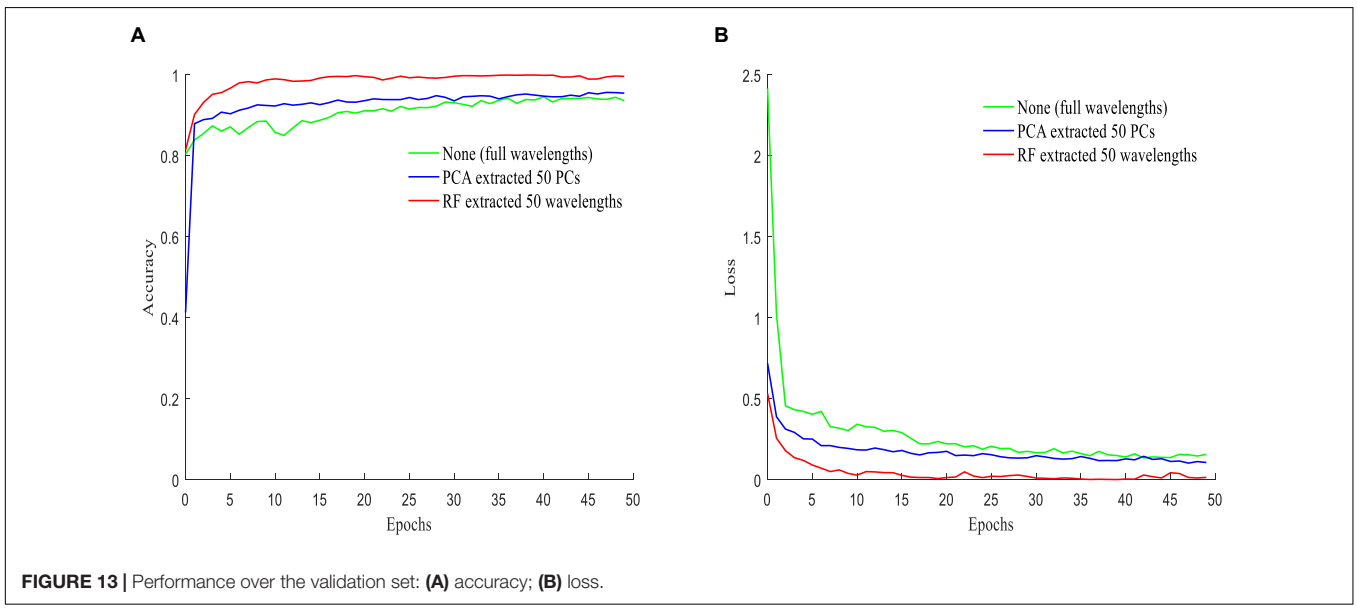


FIGURE 13 | Performance over the validation set: **(A)** accuracy; **(B)** loss.

sensitivity varies in the SDR settings. Meanwhile, we noticed that when the 50 PCs were treated as the input and the SDR was set to 2, the precision of detecting asymptomatic samples reached at 91.1620%, even higher than the performance of symptomatic detection.

By summarizing the results in **Tables 3**, **4** it can be concluded that compared with using the 50 PCs extracted by PCA as the input, using the 50 wavelengths extracted by RF may generally achieve better performance, no matter how the SDR was set. Meanwhile, due to the feature information extracted by RF and PCA is different, the SDC-3DCNN model requires specific SDR settings.

For further exploring the effect of SDR settings, we presented the detection results in **Figure 14**. The red, green, and blue pixels represent healthy, asymptomatic, and symptomatic pixels, respectively. It can be seen that when the SDR equals 1–4, the SDC-3DCNN model failed to classify certain pixels, especially the ones near the boundary between asymptomatic and symptomatic areas. This can be attributed to the over-extraction of the spectral information. When SDR equals 5, the classification result was fairly good, as more pixels can be classified correctly.

Interpretation of Feature Importance Analysis

Figure 15 shows the saliency scores of each wavelength channel, which guides us to extract the important channels for detecting healthy, asymptomatic, and symptomatic pixels from infected rice leaves.

The contributions of different wavelengths are depicted as the gray area shown in **Figure 15**. Here the spectral saliency scores were min-max normalized into the range of [0, 1]. It can be seen that all channels did not contribute equally to the classification result. The regions between 530–580 and 680–710 nm contributed the most, which were 547.2, 534.5, 551.4, 564.1, 697.4, 530.3, 695.2, 543.0, 538.7, and 568.4 nm. As shown in **Figure 16**, when compared with the top 10 characteristic wavelengths extracted by RF, it can be found that eight characteristic wavelengths overlapped with the saliency sensitive wavelengths. Through analysis of interpretation, the accuracy of the extracted wavelengths by RF is verified. At the same time, the saliency score can also be used to determine the characteristic wavelengths and reduce the dimensionality of the spectrum, thereby improving the computational efficiency.

TABLE 3 | Performance of spectral dilated convolution 3-dimensional convolutional neural network models under different SDRs with RF extracted feature inputs.

Class	Evaluating indicator	SDR				
		1	2	3	4	5
Healthy	Precision/%	94.7901	98.8107	97.3088	97.3327	97.7513
	Recall/%	99.4036	96.1350	98.3419	98.3777	98.0079
	F1-score/%	97.0420	97.4545	97.8226	97.8524	97.8794
Asymptomatic	Precision/%	92.5733	90.5362	96.8478	91.1809	94.2804
	Recall/%	91.4542	95.3800	86.4153	94.5302	91.7415
	F1-score/%	92.0104	92.8950	91.3346	92.7431	92.9936
Symptomatic	Precision/%	96.9153	96.5538	90.3685	97.1809	94.3126
	Recall/%	93.4581	94.0855	98.8845	92.5285	96.5373
	F1-score/%	95.1553	95.3037	94.4349	94.7976	95.4120
Average precision/%		94.7596	95.3002	94.8417	95.2315	**95.4481**
Average recall/%		94.7720	95.2002	94.5472	95.1455	**95.4289**
Average F1-score/%		94.7359	95.2177	94.5307	95.1310	**95.4283**
Accuracy/%		94.7640	95.1902	94.5944	95.1231	**95.4427**

The bold values denote the achieved best performance.

TABLE 4 | Performance of spectral dilated convolution 3-dimensional convolutional neural network models under different SDRs with PCA extracted feature inputs.

Class	Evaluating indicator	SDR				
		1	2	3	4	5
Healthy	Precision/%	96.4727	96.8532	96.6517	96.9132	96.5840
	Recall/%	97.8767	97.6619	97.4472	97.3756	97.1371
	F1-score/%	97.1696	97.2559	97.0478	97.1439	96.8598
Asymptomatic	Precision/%	90.2768	91.1620	89.7676	88.7874	89.8821
	Recall/%	88.2346	86.6667	89.6709	89.5631	87.6122
	F1-score/%	89.2440	88.8575	89.7192	89.1736	88.7327
Symptomatic	Precision/%	92.0673	90.5007	93.2132	92.7705	91.4471
	Recall/%	92.7841	94.0971	92.5633	91.5524	93.1792
	F1-score/%	92.4243	92.2639	92.8871	92.1574	92.3050
Average precision/%		92.9389	92.8386	**93.2108**	92.8237	92.6377
Average recall/%		92.9651	92.8086	**93.2271**	92.8304	92.6428
Average F1-score/%		92.9460	92.7924	**93.2180**	92.8250	92.6325
Accuracy/%		92.9688	92.8267	**93.2252**	92.8227	92.6531

The bold values denote the achieved best performance.

This evidence also provides some explanations for the decision-making of the model, indicating that the model might use these features to distinguish pixels in different infected states.

DISCUSSION

Advantages of Using Feature Extraction and Spectral Dilated Convolution 3-Dimensional Convolutional Neural Network Classification Models

Our results suggest that the combination of deep learning and hyperspectral imaging can be a useful approach for detecting asymptomatic infections of BLB disease in rice leaves. Previous studies have demonstrated that redundancy exists in the raw hyperspectral spectrum, leading to an increase in computational complexity (Li et al., 2017). Therefore, feature extraction is necessary. In this study, RF and PCA both can extract characteristic features from the raw hyperspectral data. Additionally, the characteristic wavelengths extracted by RF may assist the SDC-3DCNN model in achieving better classification performance than the ones extracted by PCA. Compared with PCA, the characteristic wavelengths extracted by RF could be better interpreted. The advantage of interpretability enables a multispectral camera to incorporate only the most important wavelengths instead of adopting the full wavelengths for disease detection.

Feasibility of Hyperspectral in Asymptomatic Disease Detection

We find that the healthy, asymptomatic, and symptomatic pixels of rice leaves are sensitive to different wavelengths. According to Šebela et al. (2018), Zhang S. et al. (2019), the

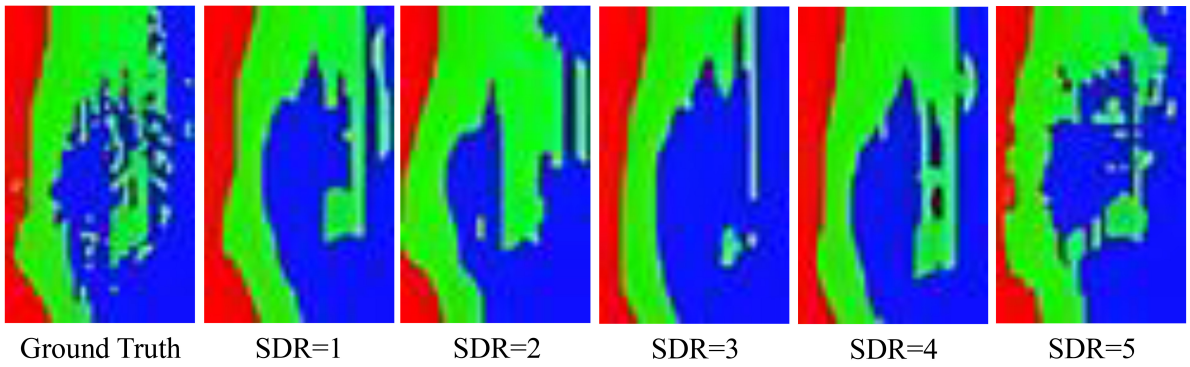


FIGURE 14 | Comparative results under different SDRs. The red, green, and blue pixels represent healthy, asymptomatic, and symptomatic pixels, respectively. The input to the spectral dilated convolution 3-dimensional convolutional neural network model is 50 characteristic wavelengths extracted by RF.

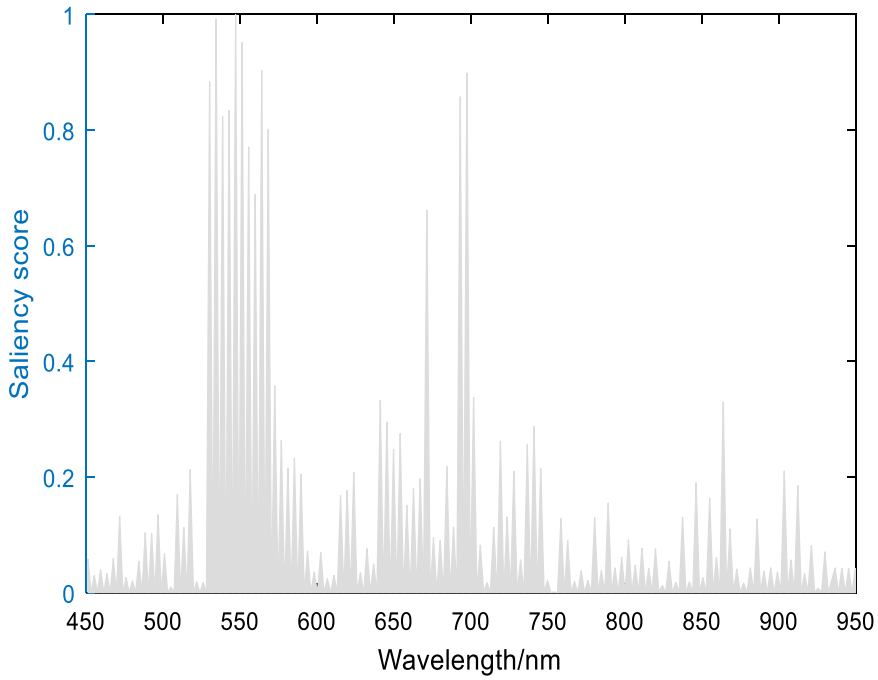


FIGURE 15 | Spectral saliency scores.

invasion of the rice BLB disease will cause inner changes such as pigment, cell structure, and leaf water content. The reflectance of different pixels is mainly affected by those inner changes in leaves, which provides a theoretical basis for disease detection based on hyperspectral imaging. In recent studies (Su et al., 2018; Wang et al., 2019; Savian et al., 2020), it has been demonstrated that accurate extraction of leaf characteristic wavelengths at different scales is feasible and facilitates the possibility of detecting the subtle spectral variation induced by the BLB disease infection. Compared with RGB images, hyperspectral imaging can obtain not only the texture and color information, but also the spectral information of the internal changes in rice leaves. This early detection can guide growers to prevent and apply pesticides in advance, thereby avoiding the BLB disease outbreak. On the other hand, the asymptomatic detection method proposed in this paper for the BLB disease can also be applied to the detection of asymptomatic infection of other plant diseases.

Potential of Spectral Dilated Convolution for Hyperspectral Imaging

At present, dilated convolution is mostly applied to increase the receptive field without pooling in 2D computer vision tasks. For the same feature map, a larger receptive field can decrease the computation complexity and improve the effect of small object recognition. Whereas the hyperspectral image is a cube, which not only includes 2D spatial information but also includes

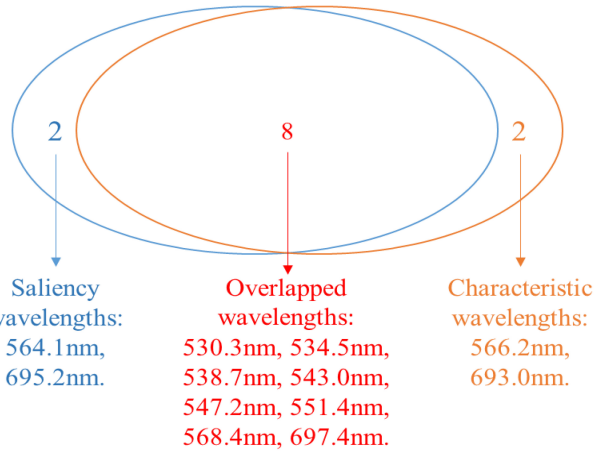


FIGURE 16 | Venn diagram for the characteristic wavelengths and the saliency wavelength channels.

the spectral information, 3D spectral dilated convolution is needed for hyperspectral image processing. There is rarely research on 3D spectral dilated convolution for hyperspectral data analysis. Xu et al. (2021) used 3D dilated convolution in hyperspectral image classification and achieved a good result. However, only spectral dimension dilated convolution was considered in the research. Considering the 3D characteristics of hyperspectral images and the information redundancy of the spectral dimension, we tested the effectiveness of 3D SDC in hyperspectral data processing. Different SDRs were compared and the SDC-3DCNN model achieved the best performance when the SDR was 5. The 3D SDC increases the receptive field not only in the spectral dimension but also in the spatial dimension, thus obtaining richer features.

Interpretation of Characteristic Wavelengths

At current, researches are paying attention to the interpretability of deep learning models because the explainable scale network can more precisely analyze the relationship between the learned scale features and different classes (Shi et al., 2022). Human-understandable results are more acceptable than undetectable black-box results, especially in the practical application of disease detection. On the basis of spectral saliency, we can find that the important wavelengths for classifying healthy, asymptomatic, and symptomatic pixels were located at 530-710 nm, which is consistent with the research results of Wang et al. (2019). In this paper, we used the saliency score to interpret the results of the model and infer the wavelengths that contributed the most to the output results. The top 10 saliency wavelengths overlapped with 8 wavelengths from the top 10 characteristic wavelengths extracted by RF, which indicated the validity of RF in extracting characteristic wavelengths. On the other hand, the saliency map method can figure out the important wavelengths, which are not extracted by RF. In addition to the significant saliency score, other interpretation methods can be further developed.

CONCLUSION

This is the first study to use hyperspectral imaging and deep learning to detect the infection of rice leaf BLB disease, particularly in the early asymptomatic stage that RGB imaging cannot detect. In this paper, RF and PCA were used to extract features from raw hyperspectral data. The detection performance of the SDC-3DCNN model with different input features and spectral dilated ratios was tested and compared. When 50 wavelengths extracted by RF were used as the input and SDR was set to 5, the SDC-3DCNN model achieved the highest accuracy at 95.4427%. In addition, the effectiveness of extracting characteristic wavelengths was verified by saliency scores, and the wavelengths with the greatest contributions were in the range of 530–710 nm. In conclusion, the combination of deep learning and hyperspectral imaging can achieve good performance for asymptomatic rice BLB disease detection. The proposed method can further evaluate the incidence of plant diseases, providing an early disease warning for farmers to apply pesticides accurately and efficiently.

DATA AVAILABILITY STATEMENT

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

AUTHOR CONTRIBUTIONS

YC contributed to the conceptualization, methodology, software, data curation, writing original draft preparation, reviewing and editing. ZZ and HX contributed to the supervision, project administration, funding acquisition, writing - review and editing. JM-O contributed to the conceptualization and methodology. PY and JF contributed to the investigation, data curation, and resources. All authors read and approved the final manuscript.

FUNDING

This work was supported by the Startup Foundation of New Professor at Nanjing Agricultural University (Grant No. 106-804005), the National Natural Science Foundation of China (Grant No. 61806097), and the Jiangsu Agricultural Science and Technology Innovation Fund (Grant No. SCX (21)3059).

ACKNOWLEDGMENTS

We would like to thank Shichao Jin from Academy for Advanced Interdisciplinary Studies, Nanjing Agricultural University, Nanjing, Jiangsu, China for his valuable suggestions. We would also like to thank Jiaqin Fan from the College of Plant Protection, Nanjing Agricultural University, Nanjing, Jiangsu, China for providing the experimental materials.

REFERENCES

Albattah, W., Nawaz, M., Javed, A., Masood, M., and Albahli, S. (2022). A novel deep learning method for detection and classification of plant diseases. *Complex Intell. Syst.* 8, 507–524. doi: 10.1007/s40747-021-00536-1

Bauriegel, E., Giebel, A., Geyer, M., Schmidt, U., and Herppich, W. B. (2011). Early detection of *Fusarium* infection in wheat using hyper-spectral imaging. *Comput. Electron. Agric.* 75, 304–312. doi: 10.1016/j.compag.2010.12.006

Bock, C. H., Barbedo, J. G. A., Del Ponte, E. M., Bohnenkamp, D., and Mahlein, A. (2020). From visual estimates to fully automated sensor-based measurements of plant disease severity: status and challenges for improving accuracy. *Phytopathol. Res.* 2:9. doi: 10.1186/s42483-020-00049-8

Cabrera Ardila, C. E., Alberto Ramirez, L., and Prieto Ortiz, F. A. (2020). Spectral analysis for the early detection of anthracnose in fruits of Sugar Mango (*Mangifera indica*). *Comput. Electron. Agric.* 105357, 1–17. doi: 10.1016/j.compag.2020.105357

Cao, F., and Guo, W. (2020). Deep hybrid dilated residual networks for hyperspectral image classification. *Neurocomputing* 384, 170–181. doi: 10.1016/j.neucom.2019.11.092

Chen, T., Yang, W., Zhang, H., Zhu, B., Zeng, R., Wang, X., et al. (2020). Early detection of bacterial wilt in peanut plants through leaf-level hyperspectral and unmanned aerial vehicle data. *Comput. Electron. Agric.* 177:105708. doi: 10.1016/j.compag.2020.105708

Chu, H., Zhang, C., Wang, M., Gouda, M., Wei, X., He, Y., et al. (2022). Hyperspectral imaging with shallow convolutional neural networks (SCNN) predicts the early herbicide stress in wheat cultivars. *J. Hazard. Mater.* 421:126706. doi: 10.1016/j.jhazmat.2021.126706

Deng, X., Huang, Z., Zheng, Z., Lan, Y., and Dai, F. (2019). Field detection and classification of citrus huanglongbing based on hyperspectral reflectance. *Comput. Electron. Agric.* 167:105006. doi: 10.1016/j.compag.2019.105006

Franceschini, M. H. D., Bartholomeus, H., van Apeldoorn, D. F., Suomalainen, J., and Kooistra, L. (2019). Feasibility of unmanned aerial vehicle optical imagery for early detection and severity assessment of late blight in potato. *Remote Sens.* 11, 21–34. doi: 10.3390/rs11030224

Hennessy, A., Clarke, K., and Lewis, M. (2020). Hyperspectral classification of plants: a review of waveband selection generalisability. *Remote Sens.* 12:113. doi: 10.3390/rs12010113

Hidayat, U. R., Nadeem, J. C., SanaUllah, M., Fahad, N., Muhammad, Y. S., and Alam Khan, R. (2017). A comparative analysis of machine learning approaches for plant disease identification. *Adv. Life Sci.* 4, 120–126.

Hsieh, T., and Kiang, J. (2020). Comparison of CNN algorithms on hyperspectral image classification in agricultural lands. *Sensors* 20, 1734–1751. doi: 10.3390/s20061734

Ji, S., Xu, W., Yang, M., and Yu, K. (2013). 3D convolutional neural networks for human action recognition. *IEEE Trans. Pattern Anal. Machine Intell.* 1, 221–231.

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

Kaitlin, M. G., Philip, A. T., Adam, C., Ittai, H., John, J. C., Larson, E. R., et al. (2020). Hyperspectral measurements enable Pre-Symptomatic detection and differentiation of contrasting physiological effects of late blight and early blight in potato. *Remote Sens.* 12, 315–330.

Koushik, N., Sarah, J., Asheesh, K. S., Soumik, S., Arti, S., and Ganapathysubramanian, B. (2019). Plant disease identification using explainable 3D deep learning on hyperspectral images. *Plant Methods* 15, 263–275. doi: 10.1186/s13007-019-0479-8

Li, Y., Zhang, H., and Shen, Q. (2017). Spectral–spatial classification of hyperspectral imagery with 3D convolutional neural network. *Remote Sens.* 9:67. doi: 10.3390/rs9010067

Lowe, A., Harrison, N., and French, A. P. (2017). Hyperspectral image analysis techniques for the detection and classification of the early onset of plant disease and stress. *Plant Methods* 13, 213–229. doi: 10.1186/s13007-017-0233-z

Lu, Y., Yi, S., Zeng, N., Liu, Y., 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

Ma, T., Jia, D., Zhou, H., Xue, Y., and Cao, J. (2018). Feature selection using forest optimization algorithm based on contribution degree. *Intell. Data Anal.* 22, 1189–1207. doi: 10.3233/IDA-173636

Mohanty, S. P., Hughes, D. P., and Salathé, M. (2016). Using deep learning for image-based plant disease detection. *Front. Plant Sci.* 7:1419. doi: 10.3389/fpls.2016.01419

Nagasubramanian, K., Jones, S., Singh, A. K., Singh, A., Ganapathysubramanian, B., and Sarkar, S. (2018). Explaining hyperspectral imaging based plant disease identification: 3D CNN and saliency maps. *ArXiv* [Preprint] doi: 10.48550/arXiv.1804.08831

Nguyen, C., Sagan, V., Maimaitiyiming, M., Maimaitijiang, M., Bhadra, S., and Kwasniewski, M. T. (2021). Early detection of plant viral disease using hyperspectral imaging and deep learning. *Sensors* 21:742. doi: 10.3390/s21030742

Pi, W., Du, J., Bi, Y., Gao, X., and Zhu, X. (2021). 3D-CNN based UAV hyperspectral imagery for grassland degradation indicator ground object classification research. *Ecol. Inform.* 62:101278. doi: 10.1016/j.ecoinf.2021.101278

Polder, G., Blok, P. M., de Villiers, H. A. C., van der Wolf, J. M., and Kamp, J. (2019). Potato virus y detection in seed potatoes using deep learning on hyperspectral images. *Front. Plant Sci.* 10:209. doi: 10.3389/fpls.2019.00209

Poona, N. K., van Niekerk, A., Nadel, R. L., and Ismail, R. (2016). Random forest (RF) wrappers for waveband selection and classification of hyperspectral data. *Appl. Spectrosc.* 70, 322–333. doi: 10.1177/0003702815620545

Qiu, J., Lu, X., Wang, X., and Hu, X. (2021). Research on rice disease identification model based on migration learning in VGG network. *IOP Conf. Ser. Earth Environ. Sci.* 680:12087. doi: 10.1088/1755-1315/680/1/012087

Rahman, C. R., Arko, P. S., Ali, M. E., Iqbal Khan, M. A., Apon, S. H., and Nowrin, F. (2020). Identification and recognition of rice diseases and pests using convolutional neural networks. *Biosyst. Eng.* 194, 112–120.

Riefolo, C., Antelmi, I., Castrignanò, A., Ruggieri, S., Galeone, C., Belmonte, A., et al. (2021). Assessment of the hyperspectral data analysis as a tool to diagnose xylella fastidiosa in the asymptomatic leaves of olive plants. *Plants* 10:683. doi: 10.3390/plants10040683

Sadeghi-Tehran, P., Virlet, N., and Hawkesford, M. J. (2021). A neural network method for classification of sunlit and shaded components of wheat canopies in the field using High-Resolution hyperspectral imagery. *Remote Sens.* 13:898. doi: 10.3390/rs13050898

Savian, F., Martini, M., Ermacora, P., Paulus, S., and Mahlein, A. (2020). Prediction of the kiwifruit decline syndrome in diseased orchards by remote sensing. *Remote Sens.* 12:2194. doi: 10.3390/rs12142194

Šebela, D., Quiï Ones, C., Cruz, C. V., Ona, I., Olejnï Čkovï, J., and Jagadish, K. S. V. (2018). Chlorophyll fluorescence and reflectance-based non-invasive quantification of blast, bacterial blight and drought stresses in rice. *Plant Cell Physiol.* 59, 30–43. doi: 10.1093/pcp/pcx144

Shi, C., Fang, L., Lv, Z., and Zhao, M. (2022). Explainable scale distillation for hyperspectral image classification. *Pattern Recognit.* 122:108316. doi: 10.1016/j.patcog.2021.108316

Shu, L., Hancke, G. P., and Abu-Mahfouz, A. M. (2021). Guest editorial: sustainable and intelligent precision agriculture. *IEEE Trans. Ind. Inform.* 17, 4318–4321. doi: 10.1109/TII.2020.3035198

Speiser, J. L., Miller, M. E., Tooze, J., and Ip, E. (2019). A comparison of random forest variable selection methods for classification prediction modeling. *Expert Syst. Appl.* 134, 93–101. doi: 10.1016/j.eswa.2019.05.028

Su, J., Liu, C., Coombes, M., Hu, X., Wang, C., Xu, X., et al. (2018). Wheat yellow rust monitoring by learning from multispectral UAV aerial imagery. *Comput. Electroni. Agric.* 155, 157–166. doi: 10.1016/j.compag.2018.10.017

Sun, Y., Wei, K., Liu, Q., Pan, L., and Tu, K. (2018). Classification and discrimination of different fungal diseases of three infection levels on peaches using hyperspectral reflectance imaging analysis. *Sensors* 18:1295. doi: 10.3390/s18041295

Tan, K., Wang, H., Chen, L., Du, Q., Du, P., and Pan, C. (2020). Estimation of the spatial distribution of heavy metal in agricultural soils using airborne hyperspectral imaging and random forest. *J. Hazard. Mater.* 382:120987. doi: 10.1016/j.jhazmat.2019.120987

Tian, L., Xue, B., Wang, Z., Li, D., Yao, X., Cao, Q., et al. (2021). Spectroscopic detection of rice leaf blast infection from asymptomatic to mild stages with integrated machine learning and feature selection. *Remote Sens. Environ.* 112350, 1–16. doi: 10.1016/j.rse.2021.112350

Wang, A., Wang, Y., and Chen, Y. (2019). Hyperspectral image classification based on convolutional neural network and random forest. *Remote Sens. Lett.* 10, 1086–1094. doi: 10.1080/2150704X.2019.1649736

Wang, C., Liu, B., Liu, L., Zhu, Y., Hou, J., Liu, P., et al. (2021). A review of deep learning used in the hyperspectral image analysis for agriculture. *Artif. Intell. Rev.* 54, 5205–5253. doi: 10.1007/s10462-021-10018-y

Wu, N., Jiang, H., Bao, Y., Zhang, C., Zhang, J., Song, W., et al. (2020). Practicability investigation of using near-infrared hyperspectral imaging to detect rice kernels infected with rice false smut in different conditions. *Sens. Actuators* 308, 113–128.

Xiao, Z., Yin, K., Geng, L., Wu, J., Zhang, F., and Liu, Y. (2022). Pest identification via hyperspectral image and deep learning. *Signal Image Video Process.* 16, 873–880. doi: 10.1007/s11760-021-02029-7

Xu, H., Yao, W., Cheng, L., and Li, B. (2021). Multiple spectral resolution 3D convolutional neural network for hyperspectral image classification. *Remote Sens.* 13:1248. doi: 10.3390/rs13071248

Xue, Y., Jiang, P., Neri, F., and Liang, J. (2021a). A Multiobjective evolutionary approach based on graph-in-graph for neural architecture search of convolutional neural networks. *Int. J. Neural Syst.* 31, 12–25. doi: 10.1142/S0129065721500350

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

Zarco-Tejada, P. J., Camino, C., Beck, P. S. A., Calderon, R., Hornero, A., Hernández-Clemente, R., et al. (2018). Pre-visual symptoms of Xylella fastidiosa infection revealed in spectral plant-trait alterations. *Nat. Plants* 1, 1–19. doi: 10.1038/s41477-018-0189-7

Zhang, B., Li, J., Fan, S., Huang, W., Zhao, C., Liu, C., et al. (2015). Hyperspectral imaging combined with multivariate analysis and band math for detection of common defects on peaches (*Prunus persica*). *Comput. Electron. Agric.* 114, 14–24. doi: 10.1016/j.compag.2015.03.015

Zhang, J., Cheng, T., Guo, W., Xu, X., Qiao, H., Xie, Y., et al. (2021a). Leaf area index estimation model for UAV image hyperspectral data based on wavelength variable selection and machine learning methods. *Plant Methods* 17:49. doi: 10.1186/s13007-021-00750-5

Zhang, J., Tian, Y., Yan, L., Wang, B., Wang, L., Xu, J., et al. (2021b). Diagnosing the symptoms of sheath blight disease on rice stalk with an in-situ hyperspectral imaging technique. *Biosyst. Eng.* 209, 94–105.

Zhang, Y., Yu, S., Zhu, X., Ning, X., Liu, W., Wang, C., et al. (2021c). Explainable liver tumor delineation in surgical specimens using hyperspectral imaging and deep learning. *Biomed. Optics Express* 12:4510. doi: 10.1364/BOE.432654

Zhang, J., Huang, Y., Pu, R., Gonzalez-Moreno, P., Yuan, L., Wu, K., et al. (2019). Monitoring plant diseases and pests through remote sensing technology: a review. *Comput. Electron. Agric.* 165:104943. doi: 10.1016/j.compag.2019.104943

Zhang, J., Yang, Y., Feng, X., Xu, H., Chen, J., and He, Y. (2020). Identification of bacterial blight resistant rice seeds using terahertz imaging and hyperspectral imaging combined with convolutional neural network. *Front. Plant Sci.* 11:821.

Zhang, S., Zhang, S., Zhang, C., Wang, X., and Shi, Y. (2019). Cucumber leaf disease identification with global pooling dilated convolutional neural network. *Comput. Electron. Agric.* 162, 422–430. doi: 10.1016/j.compag.2019.03.012

Zhao, J., Fang, Y., Chu, G., Yan, H., Hu, L., and Huang, L. (2020). Identification of leaf-scale wheat powdery mildew (*Blumeria graminis* f. Sp. Tritici) combining hyperspectral imaging and an SVM classifier. *Plants* 936, 1–13.

Zhong, Z., Li, J., Luo, Z., and Chapman, M. (2018). Spectral-Spatial residual network for hyperspectral image classification: a 3-D deep learning framework. *IEEE Trans. Geosci. Remote Sens.* 56, 847–858. doi: 10.1109/TGRS.2017.2755542

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 19 July 2022
DOI 10.3389/fpls.2022.956778

OPEN ACCESS

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

REVIEWED BY
Muhammad Imran,
King Khalid University, Saudi Arabia
Shoaib Younas,
University of Central Punjab, Pakistan
Shaomeng Shen,
The Chinese University of Hong Kong, China

*CORRESPONDENCE
Xianghu Tang
tangxh2011@iim.ac.cn

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

RECEIVED 30 May 2022
ACCEPTED 29 June 2022
PUBLISHED 19 July 2022

CITATION
Shi G, Shen X, Ren H, Rao Y, Weng S and Tang X (2022) Kernel principal component analysis and differential non-linear feature extraction of pesticide residues on fruit surface based on surface-enhanced Raman spectroscopy.
Front. Plant Sci. 13:956778.
doi: 10.3389/fpls.2022.956778

Kernel principal component analysis and differential non-linear feature extraction of pesticide residues on fruit surface based on surface-enhanced Raman spectroscopy

Guolong Shi[1,2], Xinyi Shen[1], Huan Ren[1], Yuan Rao[1,3], Shizhuang Weng[4] and Xianghu Tang[1,5]*

[1]School of Information and Computer, Anhui Agricultural University, Hefei, China, [2]School of Electrical Engineering and Automation, Wuhan University, Wuhan, China, [3]Key Laboratory of Agricultural Sensors, Ministry of Agriculture and Rural Affairs, Hefei, China, [4]National Engineering Research Center for Agro-Ecological Big Data Analysis and Application, Anhui University, Hefei, China, [5]Institute of Solid State Physics, Hefei Institutes of Physical Science (HFIPS), Chinese Academy of Sciences, Hefei, China

Surface-enhanced Raman spectroscopy (SERS) has attracted much attention because of its high sensitivity, high speed, and simple sample processing, and has great potential for application in the field of pesticide residue detection. However, SERS is susceptible to the influence of a complex detection environment in the detection of pesticide residues on the surface of fruits, facing problems such as interference from the spectral peaks of detected impurities, unclear dimension of effective correlation data, and poor linearity of sensing signals. In this work, the enhanced raw data of the pesticide thiram residues on the fruit surface using gold nanoparticle (Au-NPs) solution are formed into the raw data set of Raman signal in the IoT environment of Raman spectroscopy principal component detection. Considering the non-linear characteristics of sensing data, this work adopts kernel principal component analysis (KPCA) including radial basis function (RBF) to extract the main features for the spectra in the ranges of 653∼683 cm^{-1}, 705∼728 cm^{-1}, and 847∼872 cm^{-1}, and discusses the effects of different kernel function widths (σ) to construct a qualitative analysis of pesticide residues based on SERS spectral data model, so that the SERS spectral data produce more useful dimensionality reduction with minimal loss, higher mean squared error for cross-validation in non-linear scenarios, and effectively weaken the interference features of detecting impurity spectral peaks, unclear dimensionality of

effective correlation data, and poor linearity of sensing signals, reflecting better extraction effects than conventional principal component analysis (PCA) models.

KEYWORDS

surface-enhanced Raman spectroscopy, kernel principal component analysis, fruit pesticide residues, radial basis function, non-linear signal processing

Introduction

As a fungicide, thiram can effectively control apple scab and tomato rot, and is widely used in the cultivation of fruits and vegetables (Wang et al., 2019; Hussain et al., 2020; Gedam et al., 2021; Mbaye et al., 2022). Although the toxicity of formazan is relatively low, studies have shown that there are multiple potential harms from exposure to formazan. Currently, methods for the detection of agrochemical pollutants in fruits and vegetables include gas chromatography (Girard et al., 2021), high-performance liquid chromatography (Wei et al., 2021), gas chromatography-mass spectrometry (Ghatak et al., 2018), and liquid chromatography-mass spectrometry (Ye et al., 2020). Although these analytical techniques have good sensitivity for the quantitative detection of chemical pollutants, they still have shortcomings such as the inability to real-time monitoring, complicated operations, and cumbersome sampling process (Bereli et al., 2021). Therefore, it is necessary to propose simple and reliable methods to rapidly assess and detect pesticide residues on fruit surfaces. SERS is often used as a promising spectroscopic tool due to its advantages of high sensitivity, good specificity, simple preprocessing, and rapid spectral measurement (de Goes et al., 2019). At present, SERS has a wide range of analysis and applications and is often used to identify and detect chemical and biological species, as well as molecular imaging and monitoring at the cellular, tissue, and animal levels. It also has broad application prospects in the field of food safety (Abasi et al., 2020). Generally, SERS technology is a combination of Raman spectroscopy and nanoscience (Yoo et al., 2021), in which the molecules to be detected are adsorbed on or near the rough surface of transition metals, thereby increasing the Raman signal intensity in the local optical nanostructure region by several orders of magnitude. The effect is caused by the surface plasmon resonance of nanoparticles (Huang et al., 2020; Lin et al., 2020). As one of the most commonly used metal systems, gold nanoparticles (Au-NPs) are mainly used for SERS sensing (Zhang et al., 2017; Dowgiallo and Guenther, 2019; Szekeres and Kneipp, 2019).

Surface-enhanced Raman spectroscopy technology has great potential in detecting pesticide residues, but it still faces the following difficulties. First, good detection conditions are the basis for sensitive detection of SERS. At present, researchers have prepared various SERS substrates, but in the SERS detection of pesticide residues, they still lack high sensitivity, good repeatability, simple preparation, and low cost, which can not only enrich pesticide molecules, but also effectively enhance the suitability of the substrate (Kuo and Chang, 2014; Shen et al., 2022). In addition, there is still a lack of systematic research on the influence mechanism of different detection environments on pesticide SERS detection. Second, according to the electromagnetic enhancement mechanism, only molecules adsorbed or close to the surface of the substrate can undergo a plasmon resonance effect under light excitation, producing the SERS enhancement effect. And some pesticide molecules of weak affinity class can only produce a weak Raman signal or even no Raman signal (Krajczewski et al., 2020). In general, the SERS detection of pesticide residues on fruit surfaces is in the development stage. The ultra-sensitive detection based on SERS is prone to interference, so it is necessary to extract the characteristic information. The model of signal processing and signal recognition is still being explored, and a unified standard has not been formed. There are still many problems worth exploring.

In this work, SERS was used for the signal detection of thiram pesticide on the fruit surface, and the detection limits were determined by a combination of KPCA and partial least squares (PLS) chemometric methods after pre-processing with averaging, smoothing and differentiation, and finally, a model for the detection of thiram pesticide residues on the fruit surface was established to achieve qualitative and quantitative detection of thiram pesticide residues on the fruit surface, providing an experimental basis for the application of SERS to the detection of pesticide residues in fruits.

Related works

Surface-enhanced Raman spectroscopy refers to the phenomenon that the molecules to be tested will be adsorbed on the surface of some solid metals (gold, silver, copper, etc.) or soil particles under the irradiation of incident light, resulting in the enhancement of their local electric field (Kim et al., 2019). The intensity of the Raman spectrum obtained at this time is 10^4-10^6 times higher than that of the ordinary Raman spectrum,

which overcomes the defects of weak intensity and low sensitivity of the ordinary Raman spectrum (Zhan et al., 2019). SERS technology has also made great progress in detecting pesticide residues on the surface of fruits. Nanomaterials widely used in SERS detection of pesticide residues include noble metal sol substrates, mainly including gold, silver, and other nanoparticles, which can significantly enhance the SERS signal intensity of the analyte adsorbed on its surface (Xu et al., 2017; Ong et al., 2020). At present, the commonly used metal sol preparation methods include the electrochemical redox method, chemical deposition method, seed method, and so on, or else adding inducers such as NaC, $NaNO_3$, and cysteamine hydrochloride to the metal sol can enrich the nanoparticles and generate a large number of hot spots to improve the enhancement effect (Krajczewski et al., 2020). Stamplecoskie et al. (2011) prepared silver nanoparticles (Ag NPs) by seed method and controlled their sizes, and detected 10^{-3} mol·L^{-1} rhodamine 6G (R6G), the results showed that the optimal size of Ag NPs was 50∼60 nm, the SERS intensity on the surface of R6G is the highest, and this method is expected to be extended to other adsorbents. Xu et al. (2017) developed a surfactant-free method to prepare popcorn-like Au-NPs for the detection of Chrysanthemum cicada on the peel surface. At present, SERS has been widely used in chemical science, biological science, safety, quality inspection, etc.

Generally, the dimension of independent variables is reduced in advance, and it is hoped that fewer features are used to express the original data, to make the constructed model simpler and the results more accurate and precise, and PCA is a widely used method. The rapid screening and identification of contaminants in food contact materials is another important approach with the help of data mining technology, among which, PCA has been widely adopted as a favorable tool for data mining (Liang et al., 2021). PCA can perform dimensionality reduction on big data so that useful information in the data can be quickly extracted and classified. At present, SERS combined with PCA has been used for the rapid detection of multiple targets such as multiple disease markers (Nargis et al., 2019), and good results have been achieved. Some scholars have used this method in combination with vector machines to propose a new method to solve the problem that the original model has a large amount of computation and a slow training speed when the data is high-dimensional. It is empirically found that the results of the new method are more accurate than methods such as neural networks.

Shin et al. (2018) demonstrated the correlation between non-small cell lung cancer (NSCLC) cell-derived exons and potential protein markers in cancer diagnosis through Raman scattering spectroscopy and PCA. Ai et al. (2018) analyzed the SERS spectrum of four different food colorants using modified PCA and identified characteristic bands. Uddin et al. (2021) proposed the use of variance accumulation for selecting top features from PCA data, from segmentally folded PCA (Seg Fol PCA) and spectrally segmented folded PCA (Seg Fol PCA) FE methods Intrinsic features are selected in the transformation space of, but the non-linear relationship between transformation features generated by the PCA-based finite element method cannot be exploited. KPCA operates on the covariance of non-linear transformations of the data, allowing a more flexible functional basis to be constructed. The basic idea of KPCA is to map the linearly no separable data in the low dimensional space to the higher dimensional space through some mapping function through the kernel function so that it can be linearly separable in the high dimensional space, and then use the relevant algorithms applicable to the linearly separable data for subsequent processing. When linear mapping may not get the desired results, KPCA has more advantages than PCA. Xin et al. (2020) used a kernel function to non-linearly map the calibrated samples to a high-dimensional space, evaluated the Raman spectral reconstruction accuracy based on the relative root mean square error, and reduced bad data and non-performing samples in the sample. Sun et al. (2019) proposed a model combining KPCA and support vector machine, which effectively eliminated the influence of noise in the spectrum. Wang et al. (2021) used the synthetic minority oversampling technique (SMOTE) to predict protein-protein interaction sites and applied KPCA to remove redundant features.

Test principle and instrument reagents

Mathematical expression of Raman signal

Many fields of physics, including plasma spectroscopy, atmospheric spectroscopy, nuclear physics, and nuclear magnetic resonance, can emit information-rich spectral lines whose contours approximate the Voigt function. The Voigt function is the result of the convolution of the Gaussian function and the Lorentzian function, and its calculation process is extremely complicated. Studies have shown that the Voigt peak function is divided into a Gaussian peak and a superposition of a Lorentz peak with the same center position and half-width, and its approximate form can be expressed as (Ejiri et al., 2021):

$$V(\nu) = \theta\alpha \exp\left[-\frac{4\ln 2(\nu-\omega)^2}{\gamma^2}\right] + (1-\theta)\alpha\frac{\gamma^2}{(\nu-\omega)^2+\gamma^2}$$

In this equation, ν is the wave number, θ is the Gauss-Lorentz coefficient, α is the peak height, ω is the center position of the peak, and γ is the half-width of the peak. The spectral signal curve is formed by the superposition of dozens or hundreds of Voigt peaks. Tracing back to the source, the mathematical analysis of the vibrational spectral signal is to use the Voigt function to mathematically describe the spectral curve. The Voigt peak function is divided into a superposition

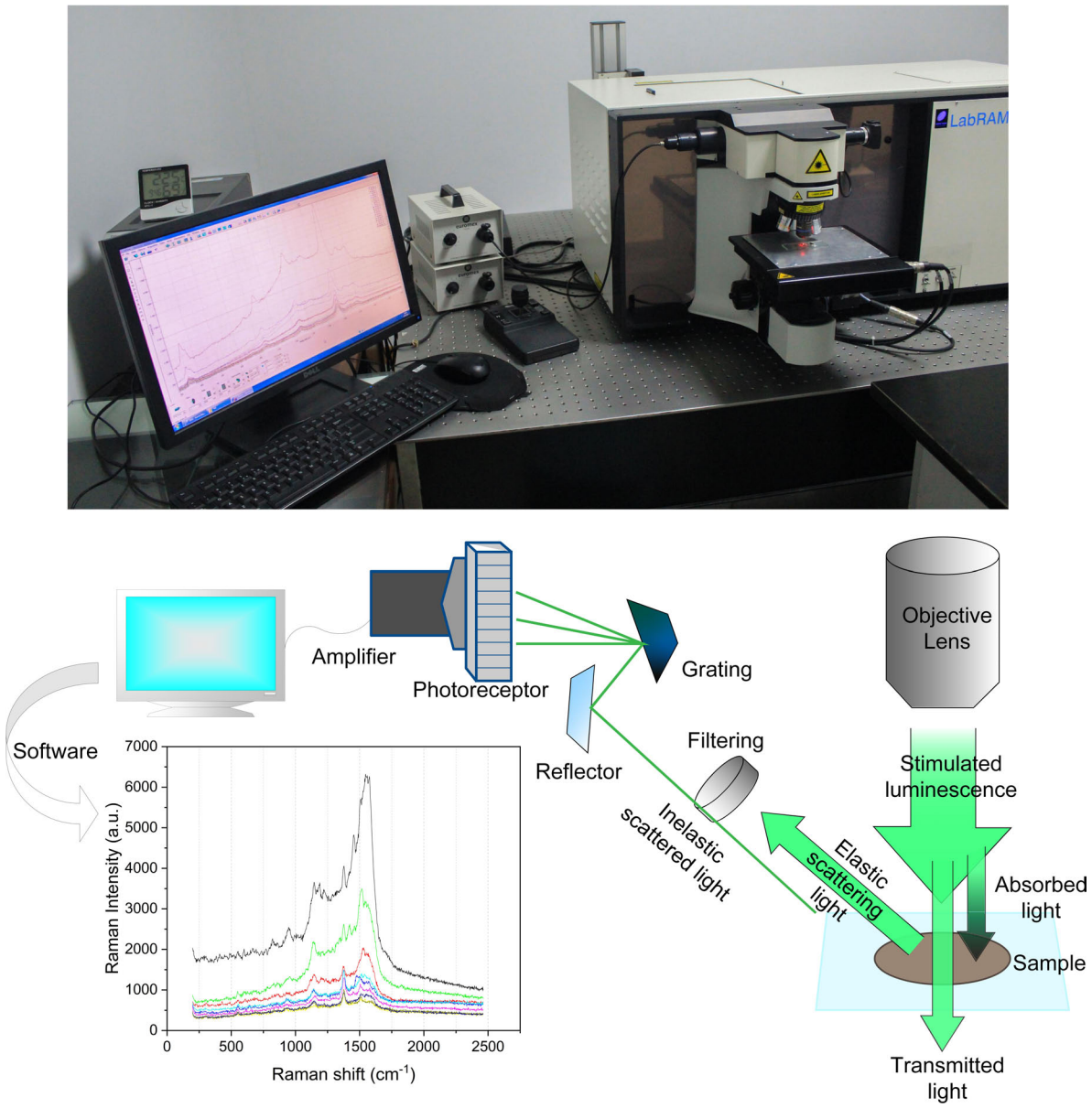


FIGURE 1
The IoT environment for the detection of principal components of pesticide residues on the surface of fruits by Raman spectroscopy.

of a Gaussian peak and a Lorentzian peak with the same center position and half-width. Therefore, the spectral peaks of the Raman and other vibrational spectra have the contour of the Lorentzian function, and its form can be expressed as:

$$L(v) = \frac{1}{\pi} \frac{\gamma_L^2}{(v-w)^2 + \gamma_L^2}$$

In this equation, v is the wave number, γ_L is the half-width of the Lorentz peak, and ω is the center position of the peak. However, the spectrum is often affected by a variety of factors, such as altitude, air pressure, or the power distribution of the laser, and the Lorentz peak profile changes accordingly. Mathematically, the effects of these factors are generally approximated by the convolution of a Gaussian function. The following is the representation of the Gaussian function:

$$G(v) = \frac{1}{\gamma_G} \left(\frac{\ln 2}{\pi}\right)^{1/2} \exp\left[-\left(\frac{v-w}{\gamma_G}\right)^2 \ln 2\right]$$

Where γ_G is the half-width of the Gaussian peak.

The research shows that the Raman spectral signal obtained by the instrument is not only the real Raman spectrum but

the result of the co-convolution of the real Raman spectrum showing the Lorentzian line shape and the instrument function showing the Gaussian line shape. The half-width of the latter depends on the resolution of the Raman spectrometer, and the half-width of the actual Raman spectral peak is much larger than the resolution of the Raman spectrometer. Therefore, if equation of G(v) is used to fit the Voigt line shape, a higher fitting accuracy can be obtained, which is suitable for various quantitative analysis situations.

Raman spectrum testing instrument

The IoT environment for the detection of principal components of pesticide residues on fruit surfaces is shown in **Figure 1**. The Raman spectrometer used in the experiment is LabRAM-HR800 from HORIBA Jobin Yvon, France, and the specification model is Horiba Jobin-Yvon LabRAM-HR800. HR800 laser confocal Raman spectrometer has the function of *in situ* spectral research, overcomes the limitation that the original Raman spectrometer can only perform offline structural analysis of materials in an indoor open environment, and realizes non-destructive and non-invasive *in situ* measurement, which provides a reliable experimental technique for real-time monitoring of physical and chemical changes of substances under specific temperature, pressure, and atmosphere (Tang et al., 2015). The Raman spectrometer has a variety of laser wavelengths to choose from and can switch gratings automatically. The detection range is between 550 and 1550 cm^{-1} and the 633 nm laser used in this experiment is used as the Raman light source, dispersion system, and data processing system, which can meet the needs of data acquisition in this experiment.

The laser reaches the surface of the sample through a series of condensing lenses, mirrors, etc. In the focused state, the radiation power of the sample per unit area reaches the maximum. The laser-focused sample produces high energy and heat. Some biological samples or substances with lower melting points often need to reduce the power during testing. The six filters on the power attenuation wheel can achieve 1/2, 1/4, 1/10, 1/100, 1/1000, and 1/10000 six-gear power reduction. The dispersive system separates the Raman scattered light in space by wavelength, usually using a grating. An important parameter of the grating is the spectral resolution (R), which is a measure of the ability to separate two adjacent spectral lines at a specific wavelength (λ). That is, $R = \lambda/\triangle\lambda$. The grating focal length (F) and the grating line density (N) are important factors to determine the spectral resolution (R), R∝F·N. The larger the F and N, the higher the spectral resolution. In addition, the spectral resolution is also related to the wavelength (λ). The larger the λ, the higher the spectral resolution. The powerful data analysis function is an indispensable part of an advanced Raman spectrometer. The Labspec 5 equipped with it has conventional data acquisition and analysis functions, and its imaging technology can generate images for different features of the spectrum (peak position, peak intensity), it also supports VB scripting language, and can also be used for Active X control in third-party applications, Labspec 5 software plug-ins can enter the spectral library and search, compatible with many commercial databases.

Substrate preparation and data collection

A 1.0×10^{-4} g·ml^{-1} solution sample of thiram solution was prepared in the laboratory, and the scanning electron microscope diagram of Au-NPs is shown in **Figure 2A**. It can be observed that the appearance of Au-NPs is spherical, and the particle diameter distribution diagram given in **Figure 2B** shows that its shape is relatively uniform. The specific operation was to dissolve 0.01 g of thiram sample in 100 ml of acetone, as shown in **Figure 2C**. Raman enhanced substrate is an important part of SERS technology. Au-NPs have stable properties and can generate local surface plasmon resonance under visible light irradiation. They are widely used in the preparation of reinforced substrates. The preparation method of Au-NPs is simple, the property is stable, and the reproducibility is high. It is an excellent material to strengthen the substrate. Therefore, 70 nm Au-NPs was prepared in the laboratory as the substrate for SERS detection (Wang et al., 2021). The test samples used in the experiment were the red Fuji apples purchased in the campus supermarket with basically the same size and weight, and Dangshan pears with almost the same size and weight, simulating the pesticide spraying process in the natural environment, that is, spraying on the surfaces of the two samples, respectively. The concentration of 1.0×10^{-4} g·ml^{-1} thiram, and then wait for 10-15 min after the fruit sample surface is automatically air-dried, take the sample epidermis, and then drop 5 microliters of Au-NPs solution on the sample epidermis, as shown in **Figure 2D**.

Limit of detection (LOD) refers to the corresponding amount of three times the instrument background signal generated by the matrix blank, or the average value of the background signal generated by the matrix blank plus three times the mean standard deviation. In the qualitative analysis of trace amounts, the LOD is used as the evaluation index to measure the enhancement effect of the substrate. For quantitative analysis, it is necessary to focus on uniformity and reproducibility. The substrate sensitivity and uniformity and repeatability have not yet reached a perfect balance. Sol-based substrates can achieve better detection limits and better SERS performance.

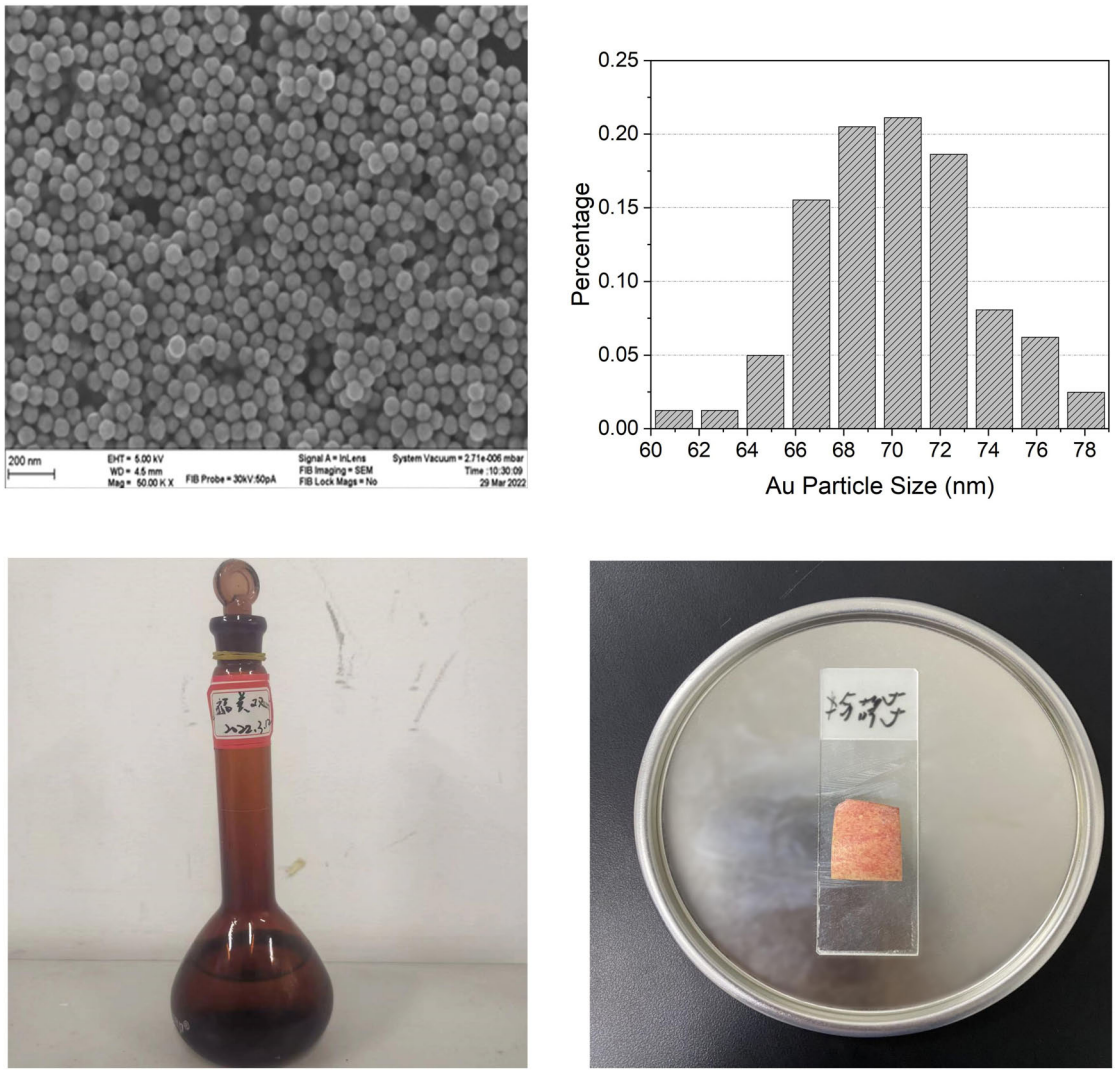


FIGURE 2
(A) Scanning electron microscope diagram of Au-NPs. **(B)** Particle diameter distribution diagram. **(C)** Preparation of thiram solution. **(D)** Dropping Au-NPs solution on the surface of the sample.

Rhodamine 6G is a kind of dye that characterizes SERS. It has strong fluorescence and has a good application effect in SERS ultra-sensitive single-molecule detection. Rhodamine groups with blocked spironolactone units can produce cation-excited fluorescence and SERS signals. Their excellent photophysical properties are widely used in fluorescent probes and SERS. To explore the uniformity and repeatability of Raman enhancement of the Au-NPs substrate used in this work, the probe molecule R6G was selected for testing, and the probe molecules located at 1510 cm^{-1} (attributed to N-H in-plane bending) and 1362 cm^{-1} (attributed to C-H in-plane bending) were tested. Statistical calculation of the SERS peak intensity values (as shown in Figure 3) at the two peaks shows that the relative standard deviation (RSD) of the two peaks is only 1.36258 and 1.63378%, indicating that the prepared Au-NPs substrates have good homogeneity and reproducibility.

During the experiment, thiram solution was sprayed on the sample for the first time and then detected by the Raman spectrum. The fluorescence signal and noise signal in the Raman spectrum experimental data obtained are very strong so that the characteristic peak signal of thiram solution is completely covered by interference, and the composition analysis of the data in the subsequent work cannot be completed. The edible wax on the skin of the fruit can not only keep it fresh but also prevent microorganisms from invading the fruit. There may be residual wax on the cleaned fruit surface, which causes a strong fluorescence signal to interfere with the Raman signal. Therefore, the experimental plan was improved in this work. The Au-NPs suspension was dropped on the tape with a pipette,

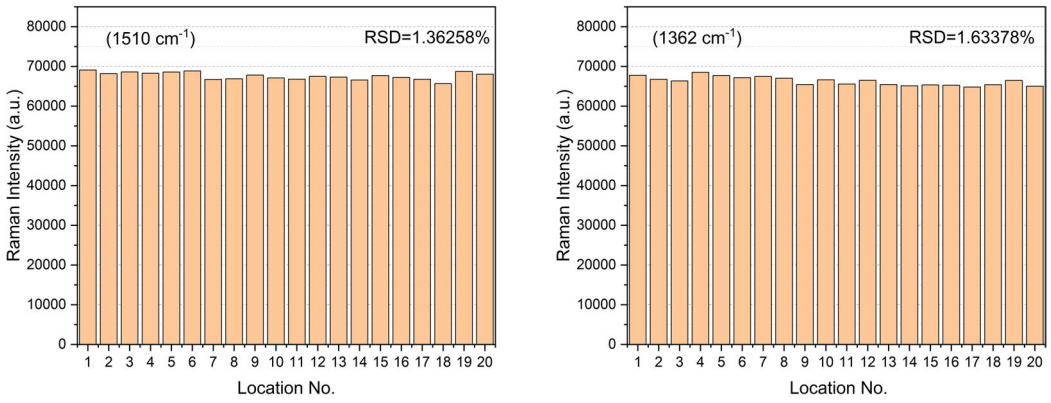


FIGURE 3
SERS spectral intensity of R6G (10^{-4} mol·L^{-1}) measured at 20 positions on Au-NPs substrate.

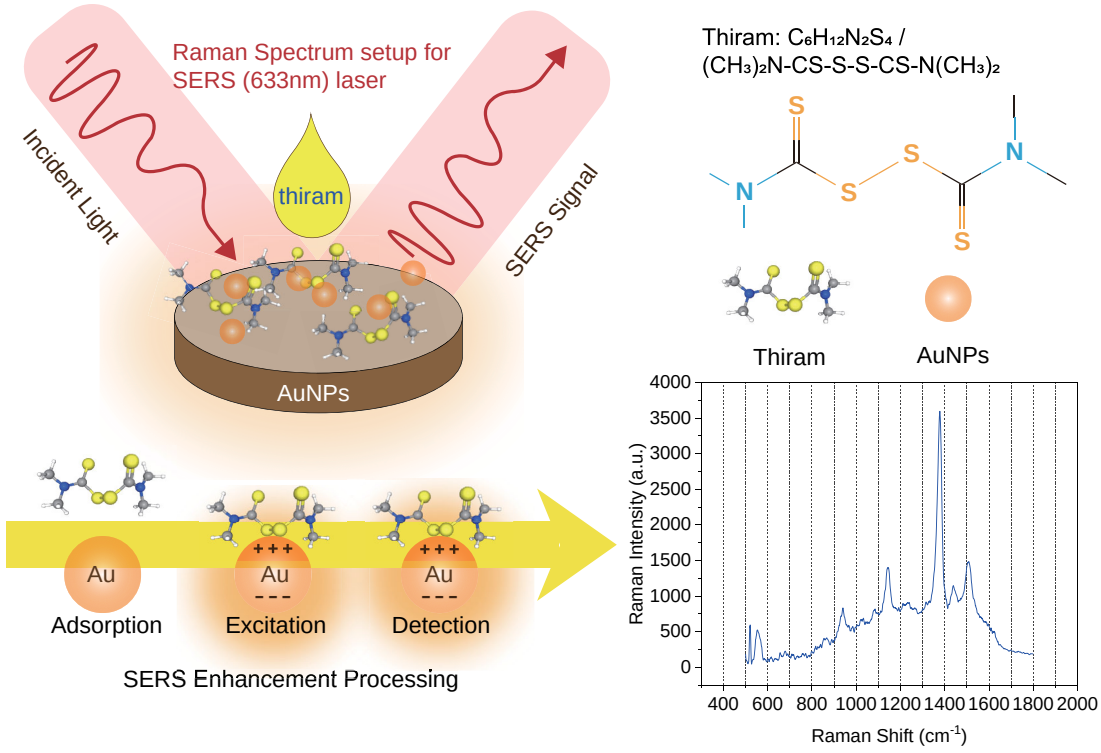


FIGURE 4
Schematic diagram of the plasma SERS model with enhancement effect with reaction time for the solution of thiram and Au-NPs.

and a drop was dropped every 8 cm or so, and it was left to stand for several hours until it was completely dry. Take different varieties of apples and pears and scrub the surface. And after drying, respectively, apply thiram solution on the surface of the fruit to air dry naturally. Adhere the tape coated with Au-NPs to the fruit containing thiram solution on the surface, peel it off after a few minutes, and place it on the Raman instrument detection table for detection (Liu et al., 2021). The transfer of pesticides to the tape can reduce the fluorescence signal and ensure the full reaction of pesticides with the substrate. Using a 633 nm light source and a 50× microscope, the LabSpec6 software collects data on different points on the surface of different varieties of apple and pear samples to obtain the raw data of the Raman spectrum on the surface of the sample. The obtained Raman spectrum data can be observed, as shown in Figure 4. The characteristic peak signal of the thiram solution indicates that the SERS can be used to obtain the original data of the sample surface. The original data includes the Raman spectra

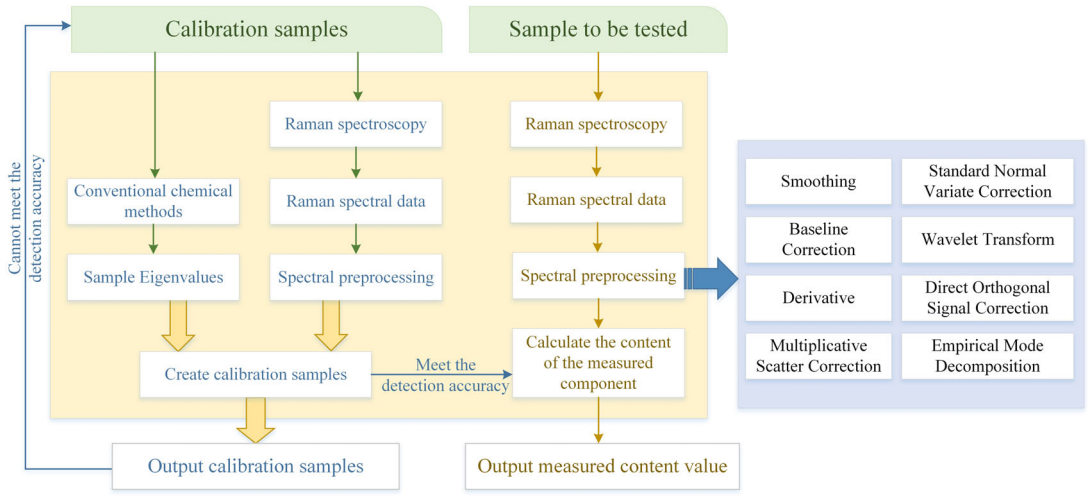


FIGURE 5
Flow chart of quantitative KPCA for SERS.

of Au-NPs on the tape, the Raman spectra of Au-NPs mixed with thiram on the fruit surface, the Raman spectra of thiram on the fruit surface, and the Raman spectra of thiram solutions with different concentrations.

Signal processing and non-linear feature extraction

Raman signal preprocessing

Since Raman scattering itself is relatively weak, the Raman spectrum is often affected by sample fluorescence, substrate fluorescence, natural light, and fluorescent light, resulting in high background and cosmic rays. When detecting the SERS signal of thiram, due to high-frequency random noise, fluorescence background, and sample unevenness, certain interference such as light scattering noise and baseline drift will be generated. The improvement of equipment often cannot eliminate these interference factors, and it is easy to affect the accuracy of subsequent prediction models. To obtain better experimental results, reduce noise, and improve the signal-to-noise ratio, the collected Raman signals must be analyzed. Perform certain preprocessing. Commonly used preprocessing methods include Smoothing, Baseline Correction, Derivative, Multiplicative Scatter Correction (MSC), and Standard Normal. Variate Correction (SNV), Wavelet Transform (WT), Direct Orthogonal Signal Correction (DOSC), and Empirical Mode Decomposition (EMD), are shown in Figure 5.

In this work, spectral averaging is used to average the SERS of thiram pesticides. Since the subsequent first-order differential and second-order differential processing will amplify the influence of noise, smoothing is used to remove the noise interference of the system and fluorescence. Improve the signal-to-noise ratio; finally, the overlapping peaks in the spectrum are separated by differential processing, and the first-order differential and second-order differential can, respectively, remove the drift that is independent of the same wavelength and linearly related. PLS is a regression modeling method of multiple dependent variables to multiple independent variables. By calculating the maximum variance between the spectral data and the target analyte, the relationship between the two is analyzed. It is suitable for complex multi-component Spectral analysis a widely used multivariate calibration method with good selectivity and predictive accuracy. PLS can eliminate the influence of data collinearity and effectively reduce the dimension of spectral data. After spectral averaging, smoothing, and differential processing, the implicit linear relationship between variables can be effectively detected due to the combination of appropriate chemometric methods, as shown in Figure 6. Therefore, the KPCA method and PLS method are used in this study to further construct the model to determine its non-linear relationship. To achieve the best fitting effect of the PLS model, the number of correction sets and prediction sets is very important. The experiment adopts the maximum-minimum strategy to establish a PLS model for samples according to a certain proportion of correction set and prediction set. First, calculate the average spectra of all candidate samples, and find the samples with the minimum and maximum distance from the average spectra to add to the calibration set. Then calculate the spectral distance between the remaining samples and each sample in the calibration set, find the samples with the minimum and maximum spectral distance from the average spectrum and add them to the calibration set, and repeat the above steps until the number of calibration sets

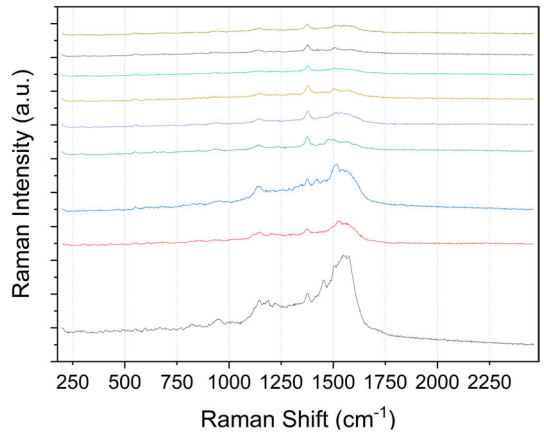


FIGURE 6
The spectrum after preprocessing the original Raman spectrum by subtracting the baseline, etc.

reaches the set value, and the remaining samples are included in the prediction set.

Non-linear kernel principal component analysis method

Surface-enhanced Raman spectroscopy spectrum is preprocessed to reduce noise interference and reduce or eliminate fluorescence background. However, because the SERS spectral data is up to thousands of dimensions and contains a lot of redundant information, the computational complexity of subsequent analysis increases, the accuracy rate decreases, and the model robustness is poor. To optimize the model and improve its prediction accuracy, the full spectrum variable modeling is usually not used, but the characteristic range spectrum is selected for analysis and processing, and the variables with a high contribution rate are extracted for modeling. Commonly used feature extraction methods include non-negative factorization (NMF), discrete cosine transform (DCT), PCA, etc. These methods obtain subject information in the sense of mathematical transformation after transforming the spectral signal.

Principal component analysis ignores the linear components with small variance and preserves the larger variance terms by processing the raw data. In this way, the dimension of effective data representation is reduced, the difficulty of problem processing is simplified, and the signal-to-noise ratio of data information is improved, to improve the prediction accuracy of the model. However, it usually requires the raw data to be Gaussian scores to extract better features, which greatly limits the practicality of this method. This is mainly because, in essence, the traditional PCA is a linear mapping method and does not do any non-linear processing, so it cannot detect the non-linear structure between the data. Therefore, many studies have extracted features between data by using non-linear PCA. On the other hand, an important feature of high-dimensional data is that the amount of data is huge, but the useful information that can be obtained from it is very limited, and there are different degrees of non-linear relationships. For this, traditional linear principal components are not sufficient.

Kernel principal component analysis uses an appropriate kernel function to project the original data space into a high-dimensional feature space. Generally, KPCA uses a non-linear kernel function to reconstruct a linear PCA, and the non-linear expansion of PCA can improve the dimensionality reduction quality of some non-linear data. KPCA maps the original data space to high-dimensional feature space and then performs PCA dimensionality reduction in the feature space.

Suppose the corresponding mapping is Φ, which is defined as $\Phi : R^d \to F, x \to \xi = \Phi(x)$. The kernel function is to implicitly realize the mapping from point x to F by mapping Φ, so that the data in the generated features satisfies the centralization condition, that is,

$$\sum_{\mu=1}^{M} \Phi(x_\mu) = 0$$

The covariance matrix in the feature space is:

$$C = \frac{1}{M} \sum_{\mu=1}^{M} \Phi(x_\mu)\Phi(x_\mu)^T$$

The eigenvalues and eigenvectors can be obtained by solving, and the test sample projection in the eigenvector space v^k is:

$$[v^k \cdot \Phi(x)] = \sum_{i=1}^{M} (\alpha_i)^k [\Phi(x_i), \Phi(x)]$$

Replacing the inner product with a kernel function, we have

$$[v^k \cdot \Phi(x)] = \sum_{i=1}^{M} (\alpha_i)^k K(x_i, x)$$

When equation the above does not hold, it needs to be adjusted

$$\Phi(x_\mu) \to \Phi(x_\mu) - \frac{1}{M} \sum_{v=1}^{M} \Phi(x_v)\mu = 1, \ldots, M$$

Then the kernel matrix can be modified as

$$K_{\mu v} \to K_{\mu v} - \frac{1}{M}\left(\sum_{w=1}^{M} K_{\mu w} + \sum_{w=1}^{M} K_{wv}\right) + \frac{1}{M^2} \sum_{w,\tau=1}^{M} K_{w\tau}$$

The KPCA algorithm essentially extracts the non-linear structure of the original data through the non-linear transformation between the data space, feature space, and category space, and combines multiple related indices into several independent comprehensive indices, to reduce the dimension of the data and solve the problem of PCA in the processing of linearly inseparable data.

TABLE 1 Predicted results of the model developed using chemometric methods.

Data	MLR	PLSR	KPCA + PLS	
	RMSECV/ ($mg \cdot L^{-1}$)	RMSECV/ ($mg \cdot L^{-1}$)	σ in KPCA	RMSECV/ ($mg \cdot L^{-1}$)
Spectra of	0.4507	0.4178	1000	3.902
653∼683,			5000	0.0347
705∼728,			8000	0.0305
847∼872 cm^{-1}			10000	0.1828

The kernel function K (kernel function) can directly obtain the inner product of the low-dimensional data mapped to the high-dimensional data, ignoring what the mapping function is, that is $K < x,y > = < \Phi(x), \Phi(y) >$, where x and y are low-dimensional input vectors, Φ is the mapping from low-dimensional to high-dimensional, and $<x, y>$ is the inner product of x and y. Kernel functions provide a link from linear to non-linear and any algorithm that can represent only the dot product between two vectors. If we first map our input data to a higher-dimensional space, the effect of operations in this high-dimensional space will be non-linear in the original space. Commonly used kernel functions are Linear Kernel (Linear Kernel) $k(x, y) = x^T y + c$, polynomial kernel(Polynomial Kernel) $k(x, y) = (ax^T y + c)^d$, Among them, the Radial Basis Function (Radial Basis Function) $k(x, y) = \exp(-\gamma||x\text{-}y||^2)$, Also called Gaussian Kernel, because it can be one of the following kernel functions:

$$k(x, y) = \exp\left(-\frac{\|x\text{-}y\|^2}{2\sigma^2}\right)$$

The radial basis function refers to a real-valued function whose value only depends on the distance of a specific point, that is,

$$\Phi(x, y) = \Phi\left(\|x\text{-}y\|\right)$$

Any function Φ that satisfies the property is called a radial vector function, Standard generally uses Euclidean distance, although other distance functions are possible. Therefore, the other two commonly used kernel functions, the power exponential kernel and the Laplacian kernel, also belong to the radial basis kernel. In this work, the SERS spectrum including radial basis function (RBF) is used to extract the main features of the spectrum in the range of 653∼683, 705∼728, and 847∼872 cm^{-1}, and the influence of different kernel function widths (σ) is discussed. Then, the support vector machine regression (SVR) algorithm was used to establish a regression model to predict the residues of thiram solution in the fruit epidermis, and the mean square error of interactive verification (RMSECV) was used to evaluate the performance of the model. The results are shown in Table 1.

It can be seen that the linear models built by multiple linear regression (MLR) and PLSR have higher RMSECV values, which may lead to lower accuracy of the prediction results; when σ is 1000, the prediction performance of the model built by KPCA combined with PLS is the worst, while the prediction performance improves when σ is 10000, but it is still weaker than when σ is 5000 and 8000. In conclusion, the model constructed by PCA combined with PLS with a σ of 8000 is the best. Its RMSECV is 0.0268 $mg \cdot L^{-1}$, the error is small, and it can accurately predict the residues of thiram solution.

The KPCA algorithm used is a qualitative and quantitative analysis model of pesticide residues written based on the measured SERS spectral data using Matlab software. The conversion equation of peak and pesticide concentration, through which qualitative and quantitative analysis of pesticide residues of unknown concentration can be carried out, and goodness of fit can be introduced to ensure that the error and accuracy of the model are within the allowable range. The goodness of fit refers to the fitting degree of the regression line to the observed value. The statistic to measure the goodness of fit is the determinate coefficient (also known as the determinate coefficient) r^2. The maximum value of r^2 is 1. The closer the value of r^2 is to 1, the better the fitting degree of the regression line to the observed value is. Conversely, the smaller the value of r^2, the worse the fitting degree of the regression line to the observed value.

Peak attribution and principal component comparison of thiram solution

According to the molecular structure and conventional Raman spectra of thiram, thiram has obvious Raman characteristic peaks at 562, 929, 1146, 1379, and 1514 cm^{-1}. The characteristic peak of 562 cm^{-1} is caused by S-S stretching vibration; the characteristic peak of 929 cm^{-1} is caused by C = S and C-N stretching vibration; the characteristic peaks of 1146 cm^{-1} and 1514 cm^{-1} can be attributed to C–N stretching and CH_3 rocking vibration; the strongest characteristic peak at 1379 cm^{-1} is caused by the C-N stretching vibration and the CH_3 symmetrical deformation vibration mode. Linear fitting was performed between the intensity (I) of the Raman peak at 1379 cm^{-1} and the concentration of the standard solution (N, $\mu g \cdot mL^{-1}$) of thiram, and the results showed that the mass concentration of thiram and the intensity of the Raman peak at 1379 cm^{-1} were linearly fitted. It has a good linear relationship. When the mass concentration range is 0.1∼5.0 $\mu g \cdot mL^{-1}$, it satisfies the linear regression equation I = 11644N + 4536.5 and the correlation coefficient $r^2 = 0.9912$.

Compared with the standard Raman spectrum of thiram solution, the characteristic peaks of the Raman spectrum of the experimental sample data obtained by MATLAB are consistent with the standard Raman spectrum of thiram in the number of characteristic peaks and Raman displacement.

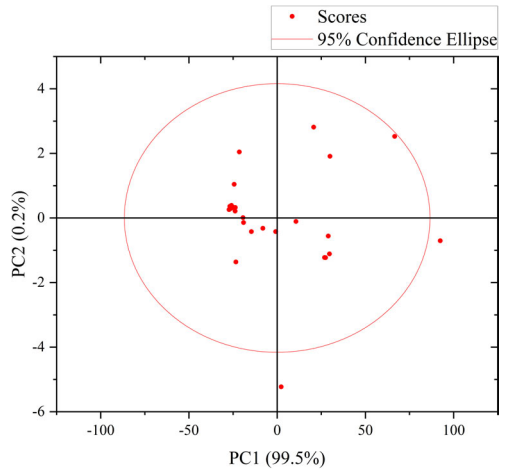


FIGURE 7
Score diagram of KPCA of thiram enhanced by Au-NPs solution.

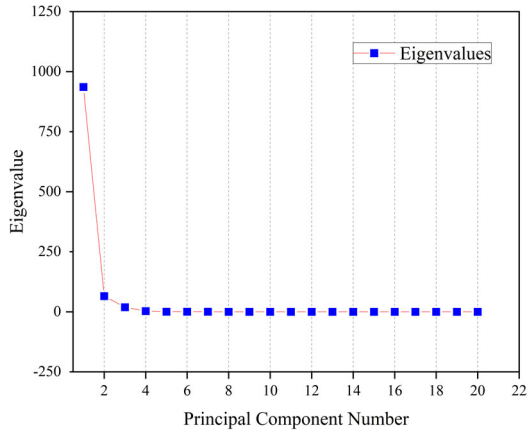


FIGURE 8
Relationship between the contribution of sample information and individual components.

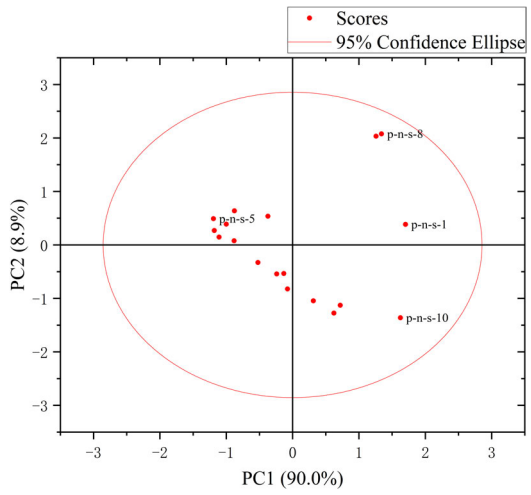


FIGURE 9
Two-dimensional scatter diagram of PC1 and PC2.

In this work, by artificially applying the standard solution of thiram pesticide to the fruit samples that were not contaminated by the pesticide residues of thiram, the residual concentrations of thiram in the fruit epidermis were 0, 0.1, 0.5, 1.5, and 10 μg/g, respectively. Its SERS was measured under a Raman microscope, and each concentration was repeated four times, and the obtained SERS was smoothed, and baseline corrected. The characteristic peaks of the SERS in the water (1:1) solution are relatively consistent, and there are characteristic peaks at 750, 830, 1165, 1560 cm^{-1}, etc., and the relationship between the intensity of the characteristic peak at 750 cm^{-1} and the concentration is the most obvious. Therefore, the characteristic peak at 750 cm^{-1} was selected to study the relationship between the peak intensity of the SERS and the concentration of fumes in the fruit epidermis. KPCA was used to determine the minimum detection limit of thiram in fruit epidermis. It can be seen from Figure 7 that the minimum detection limit of thiram in fruit epidermis is 0.1 μg/g, indicating that SERS can be used to detect thiram pesticide residues in fruit epidermis, and the minimum detection limit can reach 0.1 μ g/g.

Comparison with principal component analysis method

The visualization diagram of the 18 groups of sample data selected in this study. Through this diagram, we can compare it with the standard Raman spectrum of thiram solution after processing. From this, we can preliminarily determine whether the measured samples containing Pesticide residues are the characteristic peaks of the molecules of thiram, and with the graph of the experimental data, we can also see the similarity and difference of Raman spectra, even if the characteristic peaks of the same substance may not be the same, but the characteristics of the same substance The number of peaks is the same, and there is little difference between similar characteristic peaks. It can also be seen that the rapid detection of pesticide residues by KPCA Raman spectroscopy is more accurate.

The first two principal components PC1 and PC2 have accounted for 98.9% of the variance. It can be seen from Figure 8 that the first principal component PC1 has explained most of the variance in the sample data matrix. Figure 9 shows a 2D scatter plot generated by PC1 and PC2, where the PC2 axis is perpendicular to the PC1 axis, which is often used for data classification. In Figure 9, it can be seen that the explained variance of PC1 for the experimental sample data is 90.0%, and the explained variance of PC2 is 8.9%. And the number of principal components shown in Figure 8 is also in full agreement with this data. It can be seen that Raman

spectroscopy using KPCA combined with PLS is superior to the PCA model in terms of accuracy, precision, and stability.

Conclusion

Ultra-sensitive detection based on SERS is prone to the interference of impurities and fluorescent substances. Therefore, to play the maximum role of SERS, it is necessary to extract feature information and establish a feature recognition model, that is, the recognition module of signal processing and signal recognition model for the relevant data spectrum library. In this work, by simulating the situation of pesticide residues on the fruit surface in the natural environment and based on the SERS detection technology, the pesticide residues on the fruit epidermis were determined. It was found that the metal particles in the SERS substrate could adsorb the pesticide components in the fruit epidermis. Thus, the Raman signal is enhanced, and the interference of the fluorescent signal and noise on the surface of the fruit is prevented to a certain extent. The performance of the models processed by non-linear kernel principal components is better than that of the models processed by principal components, which proves that the former has a better dimensionality reduction effect than the latter and makes the results more accurate. The probe molecule R6G was selected for comparative testing, and the relative standard deviation (RSD) of the two peaks was statistically calculated for the SERS peak intensity value, indicating that the prepared Au-NPs substrate had an excellent enhancement effect on pesticides. Then, with Au-NPs substrate as the enhancer, the Raman peaks of the standard product of thiram solution were compared, and the characteristic peaks for qualitative discrimination of thiram solution were determined based on the assignment of spectral peaks. Using the Raman spectroscopy technique based on KPCA, the punctuation samples are standardized and preprocessed, and then the samples are non-linearly mapped by the Gaussian kernel function. Non-linear factors improve the usability and operability of measurement data and reduce computational overhead. In this method, the substances in the Raman spectrum can be classified and the pesticide residues can be detected quickly. At the same time, this work is of great value to the practical popularization of SERS.

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 was responsible for working as a supervisor for all procedures. XS was responsible for manuscript preparation and data processing. HR, SW, and YR participated in discussions and revisions. XT was responsible for providing the experimental platform and data collection. All authors contributed to the article and approved the submission.

Funding

This work was funded by China Postdoctoral Science Foundation (No. 2021M692473), Natural Science Foundation of Anhui Province (2108085QF260), the Open Research Fund of National Engineering Research Center for Agro-Ecological Big Data Analysis and Application, Anhui University (No. AE202103), and Anhui Provincial Department of Education Research Project (KJ2021A0179).

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

Abasi, S., Minaei, S., Jamshidi, B., and Fathi, D. (2020). Development of an optical smart portable instrument for fruit quality detection. *IEEE Trans. Instrum. Meas.* 70, 1–9. doi: 10.1109/TIM.2020.3011334

Ai, Y. J., Liang, P., Wu, Y. X., Dong, Q. M., Li, J. B., Bai, Y., et al. (2018). Rapid qualitative and quantitative determination of food colorants by both Raman spectra and Surface-enhanced Raman Scattering (SERS). *Food Chem.* 241, 427–433. doi: 10.1016/j.foodchem.2017.09.019

Bereli, N., Çimen, D., and Denizli, A. (2021). Optical sensor-based molecular imprinted poly (hydroxyethyl methacrylate-n-methacryloyl-(l)-histidine methyl ester) thin films for determination of Tartrazine in fruit juice. *IEEE Sens. J.* 21, 13215–13222. doi: 10.1109/JSEN.2021.3070389

de Goes, R. E., Possetti, G. R. C., Muller, M., and Fabris, J. L. (2019). Tuning of citrate-stabilized laser ablated silver nanoparticles for glyphosate detection. *IEEE Sens. J.* 20, 1843–1850. doi: 10.1109/JSEN.2019.2950161

Dowgiallo, A. M., and Guenther, D. A. (2019). Determination of the limit of detection of multiple pesticides utilizing gold nanoparticles and surface-enhanced Raman spectroscopy. *J. Agric. Food Chem.* 67, 12642–12651. doi: 10.1021/acs.jafc.9b01544

Ejiri, H., Kikuchi, R., Kumada, A., Hidaka, K., Suwa, A., Matsui, Y., et al. (2021). Excitation temperature imaging of vacuum arc based on two-line radiance method. *IEEE Trans. Plasma Sci.* 49, 1955–1961. doi: 10.1109/TPS.2021.3077971

Gedam, P. A., Thangasamy, A., Shirsat, D. V., Ghosh, S., Bhagat, K. P., Sogam, O. A., et al. (2021). Screening of onion (*Allium cepa* L.) genotypes for drought tolerance using physiological and yield based indices through multivariate analysis. *Front. Plant Sci.* 12:122. doi: 10.3389/fpls.2021.600371

Ghatak, B., Ali, S. B., Prasad, A., Ghosh, A., Sharma, P., Tudu, B., et al. (2018). Application of polymethacrylic acid imprinted quartz crystal microbalance sensor for detection of 3-Carene in mango. *IEEE Sens. J.* 18, 2697–2704. doi: 10.1109/JSEN.2018.2794607

Girard, J., Lanneau, G., Delage, L., Leroux, C., Belcour, A., Got, J., et al. (2021). Semi-quantitative targeted gas chromatography-mass spectrometry profiling supports a late side-chain reductase cycloartenol-to-cholesterol biosynthesis pathway in brown algae. *Front. Plant Sci.* 12:648426. doi: 10.3389/fpls.2021.648426

Huang, Q., Fang, C., Muhammad, M., and Yao, G. (2020). Assessment of norfloxacin degradation induced by plasma-produced ozone using surface-enhanced Raman spectroscopy. *Chemosphere* 238:124618. doi: 10.1016/j.chemosphere.2019.124618

Hussain, A., Sun, D. W., and Pu, H. (2020). Bimetallic core shelled nanoparticles (Au@ AgNPs) for rapid detection of thiram and dicyandiamide contaminants in liquid milk using SERS. *Food Chem.* 317:126429. doi: 10.1016/j.foodchem.2020.126429

Kim, J., Jang, Y., Kim, N. J., Kim, H., Yi, G. C., Shin, Y., et al. (2019). Study of chemical enhancement mechanism in non-plasmonic surface enhanced Raman spectroscopy (SERS). *Front. Chem.* 7:582. doi: 10.3389/fchem.2019.00582

Krajczewski, J., Ambroziak, R., and Kudelski, A. (2020). Substrates for surface-enhanced Raman scattering formed on nanostructured non-metallic materials: preparation and characterization. *Nanomaterials* 11:75. doi: 10.3390/nano11010075

Kuo, H. F., and Chang, C. C. (2014). Analysis of core-shell-isolated nanoparticle configurations used in the surface-enhanced Raman scattering technique. *IEEE Sens. J.* 14, 3708–3714. doi: 10.1109/JSEN.2014.2331459

Liang, H., Liu, S., Li, Z., Guo, J., and Jiang, Y. (2021). Research on infrared spectral quantitative analysis of hydrocarbon gases based on adaptive boosting classifier and PLS. *IEEE Sens. J.* 21, 20521–20529. doi: 10.1109/JSEN.2021.3096956

Lin, T., Song, Y. L., Liao, J., Liu, F., and Zeng, T. T. (2020). Applications of surface-enhanced Raman spectroscopy in detection fields. *Nanomedicine* 15, 2971–2989. doi: 10.2217/nnm-2020-0361

Liu, H., Dai, E., Xiao, R., Zhou, Z., Zhang, M., Bai, Z., et al. (2021). Development of a SERS-based lateral flow immunoassay for rapid and ultra-sensitive detection of anti-SARS-CoV-2 IgM/IgG in clinical samples. *Sensor. Actuat. B-Chem.* 329:129196. doi: 10.1016/j.snb.2020.129196

Mbaye, M., Diaw, P. A., Mbaye, O. M. A., Oturan, N., Seye, M. D. G., Trellu, C., et al. (2022). Rapid removal of fungicide thiram in aqueous medium by electro-Fenton process with Pt and BDD anodes. *Sep. Purif. Technol.* 281:119837. doi: 10.1016/j.seppur.2021.119837

Nargis, H. F., Nawaz, H., Ditta, A., Mahmood, T., Majeed, M. I., Rashid, N., et al. (2019). Raman spectroscopy of blood plasma samples from breast cancer patients at different stages. *Spectrochim. Acta A. Mol. Biomol. Spectrosc.* 222:117210. doi: 10.1016/j.saa.2019.117210

Ong, T. T., Blanch, E. W., and Jones, O. A. (2020). Surface Enhanced Raman Spectroscopy in environmental analysis, monitoring and assessment. *Sci. Total Environ.* 720:137601. doi: 10.1016/j.scitotenv.2020.137601

Shen, X., Shi, G., Zhang, Y., and Weng, S. (2022). wireless volatile organic compound detection for restricted internet of things environments based on cataluminescence sensors. *Chemosensors* 10:179. doi: 10.3390/chemosensors10050179

Shin, H., Jeong, H., Park, J., Hong, S., and Choi, Y. (2018). Correlation between cancerous exosomes and protein markers based on surface-enhanced Raman spectroscopy (SERS) and principal component analysis (PCA). *ACS sensors* 3, 2637–2643. doi: 10.1021/acssensors.8b01047

Stamplecoskie, K. G., Scaiano, J. C., Tiwari, V. S., and Anis, H. (2011). Optimal size of silver nanoparticles for surface-enhanced Raman spectroscopy. *J. Phys. Chem. C.* 115, 1403–1409. doi: 10.1021/jp106666t

Sun, H., Lv, G., Mo, J., Lv, X., Du, G., and Liu, Y. (2019). Application of KPCA combined with SVM in Raman spectral discrimination. *Optik* 184, 214–219. doi: 10.1016/j.ijleo.2019.02.126

Szekeres, G. P., and Kneipp, J. (2019). SERS probing of proteins in gold nanoparticle agglomerates. *Front. Chem.* 7:30. doi: 10.3389/fchem.2019.00030

Tang, X., Dong, R., Yang, L., and Liu, J. (2015). Fabrication of Au nanorod-coated Fe3O4 microspheres as SERS substrate for pesticide analysis by near-infrared excitation. *J. Raman Spectrosc.* 46, 470–475. doi: 10.1002/jrs.4658

Uddin, M. P., Mamun, M. A., Afjal, M. I., and Hossain, M. A. (2021). Information-theoretic feature selection with segmentation-based folded principal component analysis (PCA) for hyperspectral image classification. *Int. J. Remote Sens.* 42, 286–321. doi: 10.1080/01431161.2020.1807650

Wang, C., Wang, C., Wang, X., Wang, K., Zhu, Y., Rong, Z., et al. (2019). Magnetic SERS strip for sensitive and simultaneous detection of respiratory viruses. *ACS Appl. Mater. Inter.* 11, 19495–19505. doi: 10.1021/acsami.9b03920

Wang, X., Zhang, Y., Yu, B., Salhi, A., Chen, R., Wang, L., et al. (2021). Prediction of protein-protein interaction sites through eXtreme gradient boosting with kernel principal component analysis. *Comput. Biol. Med.* 134:104516. doi: 10.1016/j.compbiomed.2021.104516

Wei, Y., Chen, Z., Zhang, X. K., Duan, C. Q., and Pan, Q. H. (2021). Comparative analysis of glycosidic aroma compound profiling in three vitis vinifera varieties by using ultra-high-performance liquid chromatography quadrupole-time-of-flight mass spectrometry. *Front. Plant Sci.* 12:1271. doi: 10.3389/fpls.2021.694979

Xin, W. A. N. G., Zhe-ming, K. A. N. G., Long, L. I. U., and Xian-guang, F. A. N. (2020). Multi-channel Raman Spectral Reconstruction Based on Gaussian Kernel Principal Component Analysis. *Acta Photon. Sini.* 49:0330001. doi: 10.3788/gzxb20204903.0330001

Xu, M. L., Gao, Y., Han, X. X., and Zhao, B. (2017). Detection of pesticide residues in food using surface-enhanced Raman spectroscopy: a review. *J. Agric. Food Chem.* 65, 6719–6726. doi: 10.1021/acs.jafc.7b02504

Ye, D., and Wang, P., Omkar (2020). A dual-mode microwave resonator for liquid chromatography applications. *IEEE Sens. J.* 21, 1222–1228. doi: 10.1109/JSEN.2020.3018683

Yoo, J., Chang, S. K., Jung, G., Kim, K., Kim, T. S., Song, J. H., et al. (2021). Analysis of thermal characteristics of AlGaN/GaN heterostructure field-effect transistors using micro-raman spectroscopy. *J. Nanosci. Nanotechnol.* 21, 5736–5741. doi: 10.1166/jnn.2021.19491

Zhan, C., Chen, X. J., Huang, Y. F., Wu, D. Y., and Tian, Z. Q. (2019). Plasmon-mediated chemical reactions on nanostructures unveiled by surface-enhanced Raman spectroscopy. *Account. Chem. Res.* 52, 2784–2792. doi: 10.1021/acs.accounts.9b00280

Zhang, J., He, L., Zhang, X., Wang, J., Yang, L., Liu, B., et al. (2017). Colorimetric and SERS dual-readout for assaying alkaline phosphatase activity by ascorbic acid induced aggregation of Ag coated Au nanoparticles. *Sensor. Actuat. B-Chem.* 253, 839–845. doi: 10.1016/j.snb.2017.06.186

frontiers | Frontiers in Plant Science

TYPE Original Research
PUBLISHED 12 September 2022
DOI 10.3389/fpls.2022.993838

OPEN ACCESS

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

REVIEWED BY
Haider Aziz Aziz,
University of Tikrit, Iraq
Laipeng Xiao,
Tianjin University of Sport, China
Hong Chao,
Dankook University, South Korea

*CORRESPONDENCE
Jiang Zhu
jjff251314@163.com

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

RECEIVED 14 July 2022
ACCEPTED 03 August 2022
PUBLISHED 12 September 2022

CITATION
Zhu J and Sun J (2022) Ecotourism design and plant protection based on sensor network.
Front. Plant Sci. 13:993838.
doi: 10.3389/fpls.2022.993838

Ecotourism design and plant protection based on sensor network

Jiang Zhu[1,2]* and JinChun Sun[1]

[1]School of Management, Xi'an Jiaotong University, Xi'an, China, [2]Department of Design, Taiyuan Normal University, Taiyuan, China

National Forest Park is an important place for the public to carry out forest recreation activities and recognize natural habitats. With the popularization of forest tourism and the increase of forest recreational activities, the pressure on forest habitats has increased. The development of national forest parks is accompanied by opportunities and challenges. The main purpose of this paper is to analyze and study the impact of ecotourism design on plant protection based on sensor network technology. This paper analyzes the impact of tourism on the ecological environment, establishes an ecological environment monitoring system and an ecological tourism resource evaluation system, and studies the functional division of forest parks. Experimental research shows that, as a strictly protected area, the ecological conservation area basically does not conduct scenic spot development and resource mining, nor is it open to tourists. The total area is 852.92 ha, accounting for 22.31% of the total area of the forest park, allowing the ecology of the ecological conservation area to achieve sustainable and healthy development.

KEYWORDS

sensor network, ecotourism design, plant protection, forest park, conservation area

Introduction

With the development of science and technology and the leap of economy, the integration of agriculture and eco-tourism is developing rapidly, and the agro-eco-tourism industry has promoted the transformation of agricultural industrial structure and the improvement of efficiency, and has been flourishing in all parts of my country. The basic characteristics of agricultural ecotourism include dependence on ecological agriculture, taking ecological environment protection as the core, combining with rural culture, and extensive participation in tourism activities. At the same time, due to the insufficient understanding of the role of plant protection in tourism, the surrounding environment has deteriorated and the risk of large-scale outbreaks of pests and diseases has increased. From the academic point of view, ecotourism mainly focuses on planning, economy and ecology, but there are few theoretical studies on the role of plant protection on ecotourism (Awasthi and Bollas, 2020; Manoj et al., 2020). The construction of

ecological agriculture has laid a good material foundation and pollution-free security for the development of tourism. The establishment of green agricultural product bases in tourist attractions and the adoption of science and technology to plant crops are not only conducive to the promotion and application of agricultural technology, but also provide green food for tourists. Tourism and eco agriculture are combined, and they complement each other and promote each other, so as to realize the ecological, social and economic benefits of agricultural eco-tourism.

In related research, Sambo et al. (2020) mentioned that wireless underground sensor network (WUSN) faces the problem of wireless underground communication (WUC), which greatly attenuates ground signals. Wireless sensor network unit includes data acquisition unit, data transmission unit, data processing unit, and energy supply unit. At the end of it is a sensor that can sense and inspect the external world. His sensors communicate wirelessly. From this, a WUSN path loss for precision agriculture is proposed, called WUSN-PLM. To achieve this, the proposed model is based on an accurate prediction of the complex permittivity (CDC). Martin et al. (2021) proposed an algorithm to map direct normal irradiance (DNI) in thermal solar power plants using a mobile robotic sensor network (RSN). The algorithm selects measurement points and assigns RSN accordingly for dynamic estimation of DNI. Sensor network realizes three functions of data acquisition, processing and transmission. Together with communication technology and computer technology, it constitutes the three pillars of information technology. The performance of the algorithm is evaluated using a generic thermal solar power plant with a fleet as a simulated case study.

Based on sensor network technology, this paper analyzes and studies the impact of ecotourism design on plant protection. This paper first introduces the meaning of ecotourism, and analyzes the four principles of ecological planning, namely, the principle of ecological protection, the principle of overall optimization, the principle of adapting to local conditions, and the principle of landscape heterogeneity. The impact of water bodies, soil, plants, and wild animals; then, the data processing technology is analyzed, the ecological environment monitoring system and the ecotourism resource evaluation system are established, and the comprehensive ecological characteristics are analyzed. Note that For the development and ecological protection of tourism resources, combined with the comprehensive improvement of key river basins and regional environment, the environmental management of tourist areas has been strengthened, the environmental impact assessment has been carried out for the planning, development and construction projects of some tourist areas, the pollution prevention and control efforts have been strengthened, and a number of polluting enterprises in scenic tourist areas have been shut down, relocated and treated within a time limit (Palani et al., 2020).

Design research

The meaning of ecotourism

The meaning of ecotourism has been constantly changing, and the process of changing the definition of forest park is similar, but the emphasis is still different. Tourism, with distinctive ecological environment as its main landscape, takes sustainable development as the concept, takes the protection of ecological environment as the premise, takes the harmonious development between man and nature as the criterion, and relies on a good natural ecological environment and a unique human ecosystem. The key points include ecotourism tourist motivation, ecotourism resource level, socioeconomic capacity of the region, and ecological capacity of forest parks. Because ecotourism is different from other forms of tourism, it appeared in a later period, so it is not very familiar to the public, and it is easy to be confused with other concepts such as mass tourism, nature tourism, and sustainable development tourism that have appeared in research for a long time (Baghaee et al., 2020; Venkataravanappa et al., 2020).

Distinguishing analysis of similar concepts and categories:

One is the difference between ecotourism and mass tourism. Compared with ecotourism, mass tourism emerged earlier, and ecotourism is a derivative of mass tourism. Because mass tourism is an early concept and form of tourism, its connotation does not include the protection of natural environment and ecological resources, the threshold for development and introduction is low, and it is easy to develop and implement, but there are also problems of damage and negative impact on forest resources. The first meaning of mass tourism refers to that the range of participants in tourism activities has been extended to ordinary working people. The second meaning is that modern tourism activities began to form a mass tourism model represented by organized group package tourism, and formed a dominant tourism form among the general public. In terms of planning management, development goals, and stakeholders, ecotourism prioritizes resource protection and avoids the sacrifice of forest resources. Therefore, it is quite different from mass tourism.

The second is the difference between ecotourism and nature tourism. The difference between the two is small, the main reason is that both rely on natural resources and are closely related to animal and plant resources, as well as physical resources such as mountains and rivers. Natural tourism takes natural resources as the core of development power. Compared with mass tourism, it is more inclined to be close to natural resources, but it also lacks the popularization of environmental protection awareness. Therefore, tourism activities will still have an impact on natural resources. On the basis of natural tourism, ecotourism emphasizes the protection of the natural environment and ecological resources, from people-oriented to ecological resources-oriented, and realizes

the value of ecotourism through the interaction between people and ecological resources in the natural environment. The basic principles of people-oriented management include: paying attention to people's needs, encouraging employees, cultivating employees, and people-centered organization design.

Among these three tourism concepts, the origin of sustainable tourism comes from ecotourism. From the perspective of sustainable development, sustainable tourism is a persistent and continuous tourism behavior. The difference between sustainable development tourism and ecotourism is that sustainable development tourism contains more abundant concepts. There are various forms and processes of sustainable development tourism, and it is not limited to ecotourism. The main connotation of sustainable tourism includes improving people's understanding of the environmental and economic impact of tourism, strengthening people's ecological awareness, promoting the fair development of tourism, improving the quality of life in tourism reception areas, providing high-quality tourism services to tourists, and protecting the environmental quality on which future tourism development depends.

Principles of ecological planning

Principles of ecological protection

The ecological elements of the ecosystem play an important role in the stability of the resource utilization of the forest park. Therefore, in the planning, it is necessary to respect and conform to nature, and on this basis, adhere to the principle of ecological security to provide the maximum survival for the biological groups in the forest park. It is necessary to improve the ability to protect ecological resources, maintain the connectivity between various landscapes, and create a living environment space suitable for biological habitats. Ecological security has the characteristics of integrity, irreversibility and long-term. Specifically, ecological security is a state of human living environment or human ecological conditions. Ecological security is a dynamic concept. Ecological security emphasizes people-oriented, and maintaining ecological security requires costs. Focusing on the comprehensive protection of landscape resources, a targeted forest park planning scheme is formulated, and ecological protection and forest tourism are always positioned as the primary functions of the park, and development is carried out on the basis of sustainable development (Blake et al., 2021).

The overall optimization principle

The principle of overall optimization should be considered, so that the internal functional division of the forest park can be adapted to the surrounding environment, and the symbiosis and prosperity of the forest park and the surrounding environment should be ensured through overall protection and system optimization. The attributes and functional properties of each component in the park are different, but the overall style and form should be unified, and the overall characteristics and image characteristics of the forest park should be maximized through analysis, and each element must be balanced in the ecological environment. We can't ignore one and lose the other, take the overall optimization as an important organic part of ecotourism, and strive to create an ecologically harmonious forest ecosystem.

The principle of adjusting measures to local conditions

When planning and designing suburban forests, it is necessary to respect the *status quo* of the natural, cultural and economic conditions of the region, and carry out practical designs based on research and analysis of superior planning, so as to highlight the characteristics of suburban forest parks. It is necessary to focus on ecological conservation and cultural protection, and design various places for forest tours, popular science education, and outdoor recreation according to local conditions. The construction of suburban forest parks is to protect the original ecological environment and develop its unique landscape on this basis. Select native plants, give full play to climatic characteristics, properly use exotic plants and cultivate new high-quality tree species to form a multi-layer mixed structure of trees, shrubs, grasses, and ground cover. Create infrastructures such as water landscapes, buildings, and sketches that are rich in local characteristics and forest park orientation, so that various landscape spaces can cooperate with each other to form a forest tourism environment with changing scenery and increase the fun of the tour. The economic characteristics of suburban forestry include the duality of suburban forestry, the diversity of suburban forestry production, and the economy of suburban forestry. Among them, the economy of suburban forestry is reflected as follows: on the one hand, facing the vast cities, suburban forestry has fast information, wide product sales and broad market prospects; On the other hand, the suburban forestry cannot be completed like the urban garden department, and the construction funds are allocated by the state to solve the construction of suburban forestry.

The principle of landscape heterogeneity

Landscape heterogeneity refers to the variability of landscape elements in the landscape system, which is beneficial to the division of the spatial pattern of the ecological environment and to improve patch heterogeneity and complexity. The degree of landscape heterogeneity is related to the biodiversity of the landscape. In the construction, the existing landscape heterogeneity should be used and strengthened, the vertical landscape design should be enriched, the aquatic plant habitat system should be established, and the ecological balance should be maintained. Strengthen the protection of ecologically sensitive areas, provide a heterogeneous living environment for the creatures in the park,

strengthen the construction of landscape heterogeneity, and create a heterogeneous tourism space for tourists.

The impact of tourism on the ecological environment of forest parks

The impact of tourism activities on the ecological environment is usually reflected in specific locations in tourist areas. Because the locations of natural landscapes and cultural attractions that attract tourists are relatively fixed, they account for a small proportion of the total area. Among them, the impact of tourism activities on vegetation and soil is relatively concentrated, which are fixed components of the ecosystem, while the impact on water bodies and wild animals is relatively scattered (Alkhalifa and Almogren, 2020; Khodadadiangostar et al., 2020).

Impact on water bodies

The impact of tourism activities on water bodies is mainly reflected in the quality of water quality, such as: eutrophication of water bodies, discarding of domestic waste, mass reproduction of planktonic algae, and water pollution caused by discarded suspended solids. Water pollution will directly affect the death and accumulation of aquatic animals and plants. The accumulation of pollutants causes harmful gases to be produced in water bodies. The human excrement produced by the residents living in the forest park also greatly pollutes the water quality, and the hotels, shopping streets, restaurants and other tourist service places opened near the water resources will also produce non-degradable waste when entertaining tourists. It takes a long time to purify the wastes to remove the pollution. If these pollutants are not treated scientifically and properly, it will cause water pollution in the forest park and surrounding water bodies. In the surface fresh water system, phosphate is usually the limiting factor of plant growth, while in the seawater system, ammonia nitrogen and nitrate are often the limiting factors of plant growth and total production. The substances that lead to eutrophication are often the nutrients with limited content in these water systems. Accordingly, the prevention and control measures include controlling the input of exogenous nutrients and reducing the load of endogenous nutrients.

Although the water body has a self-purification function and has a certain ability to withstand external influences, the pollution caused by the short-term is not significant, but after a long period of time, when the pollutants accumulate to a certain level, the problem will become prominent. The aquatic ecosystem has a dynamic equilibrium system, and usually all kinds of organisms in the aquatic ecological environment are in the trend of mutual balance. However, when various tourism activities follow, the water ecological balance will be broken, and the linkage of the entire ecological chain will begin to fail, which will reduce the stable development of aquatic organisms and the protection of diversity, will cause more serious ecological damage.

Effects on soil and plants

Because the growth of vegetation is inseparable from the soil, tourism activities, especially the impact of trampling on the soil, will indirectly affect the growth of vegetation. Vehicles roll on the soil, causing the soil to be compacted, or tourists trampling on the lawn, causing damage to the vegetation surface, severe surface exposure and serious soil erosion, which affect the growth and vitality of trees, which in turn has a series of cyclic effects on soil and vegetation. It is also related to the location of tourist routes and scenic spots. For scenic spots that are used more frequently, the scope of influence will expand with the extension of usage time and the increase of tourists' demand for space.

The most direct consequences of tourism activities on soil and plants are: the reduction of surface vegetation coverage and the reduction of species diversity, which in the long run will affect the ecosystem of suburban forest parks.

Impact on wild animals

Tourism activities create a stressful environment for animals, and noise pollution from development and tourism activities has the greatest impact on wildlife. When animals live in an environment that exceeds the normal volume for a long time, they will become more and more timid and unresponsive. When encountering a moving vehicle or an advancing ship, they cannot avoid danger in time, resulting in an increase in accidental mortality, and even put some endangered animals on the verge of extinction. The trampling of natural ground cover by tourists and the uncontrolled picking of flowers and fruits of plants will also have a negative impact on the animal habitat in the forest park, which limits the living space and activity range of animals to a large extent. At the same time, hunting and hunting activities have a great impact on wild animals. The endless hunting activities of humans have led to a sharp decrease in the number of wild animals. The destruction of the overall ecological environment also has a certain impact on the number of wild insects (Nobar et al., 2020; Younus et al., 2020).

Algorithm research

After the forest data collection is completed through the image, according to the characteristics of the Rikola hyperspectral remote control data, on the basis of completing the initial quality assessment of the image, the hyperspectral image is reproduced in the system and process. Compared with the traditional low spectral resolution remote sensing technology, hyperspectral remote sensing provides a wider application in earth observation and environmental investigation, which is mainly reflected in the following aspects:

the resolution and recognition ability of ground objects is greatly improved, the imaging channel is greatly increased, and the influence of other interference factors is largely suppressed in the spectral space.

System calibration

In the process of acquiring hyperspectral images by the UAV platform, there will inevitably be systematic errors due to the limitations of the instrument itself and the measurement method, and such errors need to be corrected.

Vegetation hyperspectral images represent ground object information by pixel brightness value (DN value), but due to the influence of systematic errors, DN value cannot truly reflect the spectral properties of ground objects, and it is necessary to use measurement information representing specific signals obtained during the testing process. Convert the DN value of the original image to the reflectivity of ground objects, the formula is as (1).

$$\rho_t = \frac{DN_t - DN_1}{DN_2 - DN_1}(\rho_2 - \rho_1) + \rho_1 \quad (1)$$

In the formula, ρt and DN_t are the reflectivity and DN value of the original image target pixel, respectively, ρ_1 and ρ_2 are the reflectivity of different calibration oil cloths, respectively, DN_1 and DN_2 are the DN values of different calibration oil cloths, respectively. The main application of traditional imaging remote sensing technology is qualitative analysis, and the accuracy of some quantitative analysis results is not ideal. This is obviously due to the limitations of the spectrum and spatial resolution of the imaging sensor, the interference of the atmosphere and soil background, etc., which greatly suppresses the influence of other interference factors in the spectral space, which is of great help to improve the accuracy of quantitative analysis results.

Post-processing correction

The UAV images obtained by the frame-based imaging method in the flight experiment in this study are affected by factors such as the imaging principle and the environment. There are subtle differences in position and attitude among the 45-band images, resulting in the generated hyperspectral cubes with different bands in each band. Cannot completely overlap, we used the Coregister folder tool of Pix4D software for band registration. The time for a drone to fly once is about 20–30 min. Affected by changes in sunlight conditions, there will be gradient differences in the radiance between different flight zones, and the image will often show uneven color and brightness. Using equations (2) and (3) can effectively correct the irradiance to the normal level.

$$L_{jc}(\lambda)_{at\_sensor} = L_j(\lambda)^*_{at_{sensor}} C_j(\lambda) \quad (2)$$

$$C_j(\lambda) = E_j(\lambda)/E_{ref}(\lambda) \quad (3)$$

Among them: $L_{jc}(\lambda)_{at\_sensor}$ represents the image after irradiance consistency correction; $L_{jc}(\lambda)_{at\_sensor}$ represents the jth original image; $C_j(\lambda)$ represents the multiplicative correction factor of j; $E_j(\lambda)$ represents the record of j Irradiance value; $C_{ref}(\lambda)$ represents the irradiance value of the reference image. Note that image gradient can regard image as a two-dimensional discrete function, and image gradient is actually the derivation of this two-dimensional discrete function. Image edge is generally realized by gradient operation of image.

Experimental study

Data processing technology

Using ArcGIS and ENVI software to preprocess the collected basic data, the main technical tools involved are as follows:

Image mosaic

Image mosaicking is a tool for splicing multiple adjacent image maps into a large-scale image with spatial connection due to the limitation of the area of remote sensing image maps. The methods of cutting lines and feathering are often used for boundary processing. After multiple adjacent images across the state, the images need to be stitched together using the mosaic tool. The whole process of image mosaic technology includes image preprocessing, image registration, establishment of transformation model, unified coordinate transformation, fusion, and reconstruction.

Image cropping

Due to the difference between the directly acquired data space range and the actual range of the research object, when conducting regional research, image cropping is often used to extract the specified range of the original image. This paper mainly uses the mask extraction tool to extract various types of data within the research area data, so that the spatial extent of different data layers after processing is consistent.

Resampling

Resampling is the use of ArcGIS to process existing raster data to convert its resolution to a specified size. Since the delineation of ecological space involves a lot of index data, it is necessary to use the resampling tool in the data management tool to unify it into a grid with a resolution of 100 m × 100 m, and then perform overlay calculation.

Reclassification

Reclassification is to reclassify and assign values according to the required sequence according to the size of the original data values, so as to obtain a new set of data. Cover the original values of indicators, etc., and use the reclassification in the

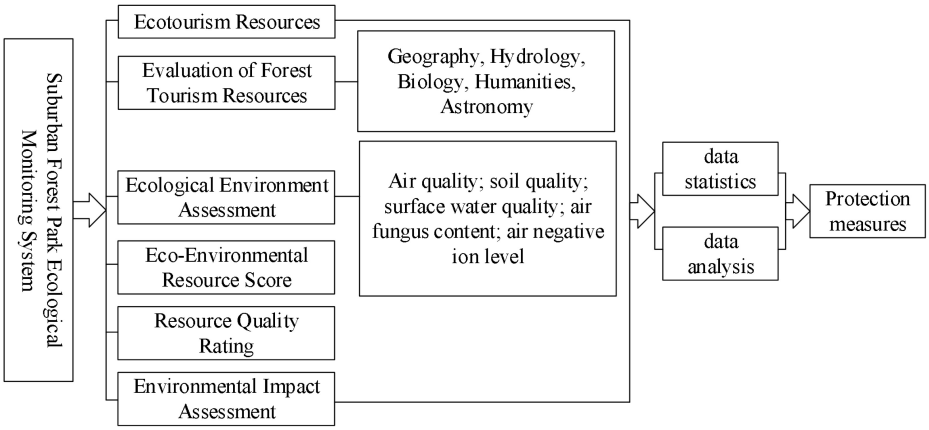


FIGURE 1
Ecological environment monitoring system diagram.

spatial analysis module to reclassify the indicator data interval and assign sensitive values.

Neighborhood analysis

In Figure 1, re-classification is performed according to the size of the original value, the data set and the original value are obtained, and the re-classification in the spatial analysis module is used to re classify the index data interval and allocate the sensitive value. Since the prior art cannot directly obtain the topographic relief related to the topographic factors, it is necessary to select focus statistics in the neighborhood analysis module of ArcGIS spatial analysis through DEM elevation data, and select 3 × 3 pixels to extract the maximum and minimum values. Raster calculation. There are many forms of data organization and expression of digital elevation model, including regular rectangular grid and irregular triangular grid, which are commonly used in land use engineering.

Grid calculation

Using the grid calculator, according to the evaluation model formula of ecological space delineation, the multi-raster layer data is substituted and the superposition operation is performed to obtain the spatial distribution relationship between the evaluation factors.

Fuzzy classification

In the evaluation of the importance of ecological service functions, the evaluation results of each subsystem need to be normalized. Select the fuzzy classification in the superposition analysis, input the original evaluation value, select the classification value type as a linear function, and assign the data membership between 0 and 1. The core of fuzzy set is the determination of membership function, which has a great impact on the application effect of fuzzy set. The process of determining the membership function is closely related to the practical application background, and there is no general method.

Interpolation analysis

Interpolation analysis is by inputting coordinate points with specific values, and using kriging interpolation or inverse distance weighting (IDW) in the interpolation analysis of the spatial analysis module to assign spatial evolution trends to point coordinate values. Due to the spatial evolution of meteorological station data, it is manifested as the spatial distribution trend of meteorological data such as rainfall and wind speed.

Ecological environment monitoring system

The construction of the ecological environment monitoring system can comprehensively understand the resources of suburban forest parks, promote the integrity of forest eco-tourism resources, enable forest tourism resources to achieve the most scientific and effective development and utilization, convert various benefits into value, and improve the forest resources. Quality and value of park ecotourism resources.

The main function of the ecological environment monitoring and evaluation system is to confirm the ecological level of the suburban forest park, reduce the adverse impact, and propose reasonable ecological protection measures. Take the ecological environment impact assessment work as a routine work and put it into the various links of the ecological tourism development and construction of suburban forest parks. The research contents mainly include: ecological environment assessment and environmental impact assessment reflecting changes in environmental quality; graded assessment of scenic resource quality and put forward a scientific and

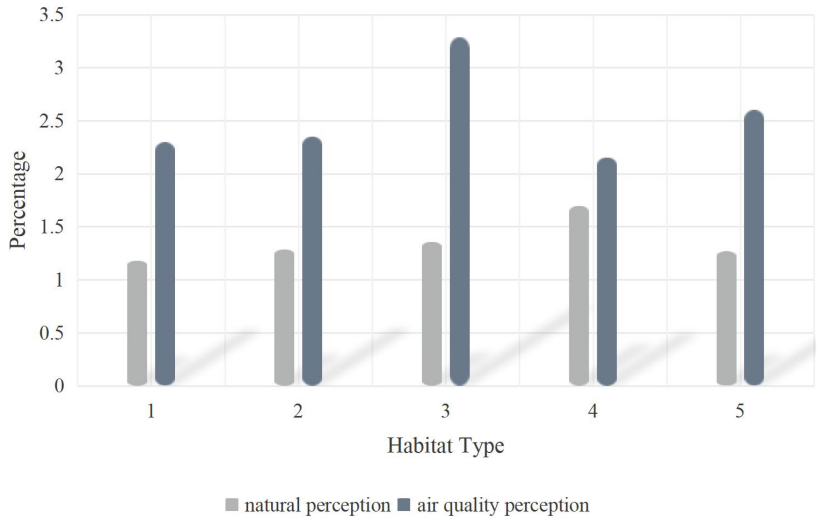


FIGURE 2
Analysis of differences in natural perception and air quality perception of different habitat types.

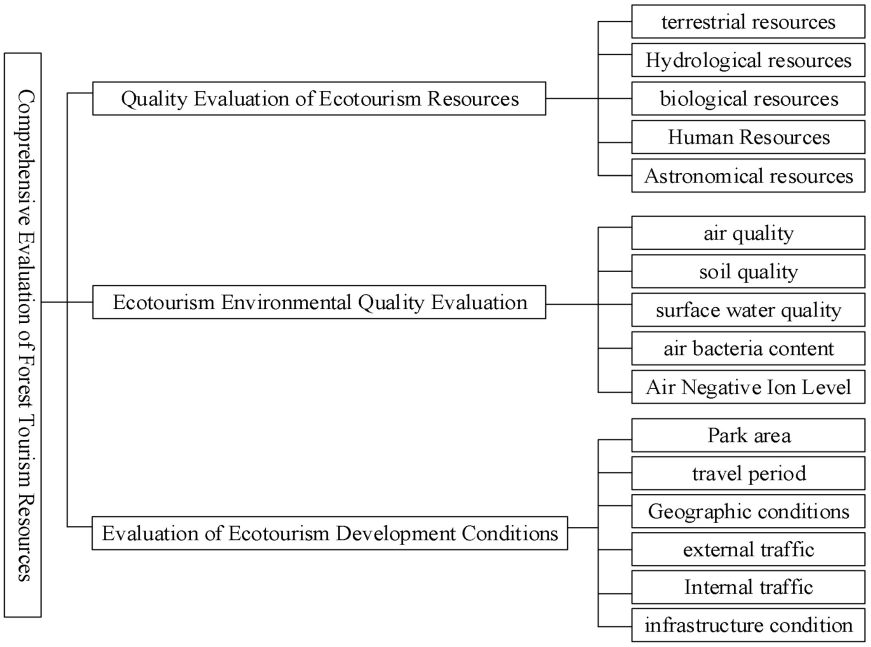


FIGURE 3
Evaluation framework of ecotourism resources.

effective resource development model. Monitoring should not only ensure systematization and timeliness, but also clearly reflect the existing problems in suburban forest parks, so that effective protection measures can be taken more quickly, so as to continuously optimize the planning and construction management of suburban forest parks, and solve the problems of tourism resources and environment protection issues.

Evaluation system of ecotourism resources

In Figure 2, different habitats are analyzed, and reasonable ecological protection measures are proposed through differences in natural perception and air quality to enhance the role of forest resources. In order to give full play to the

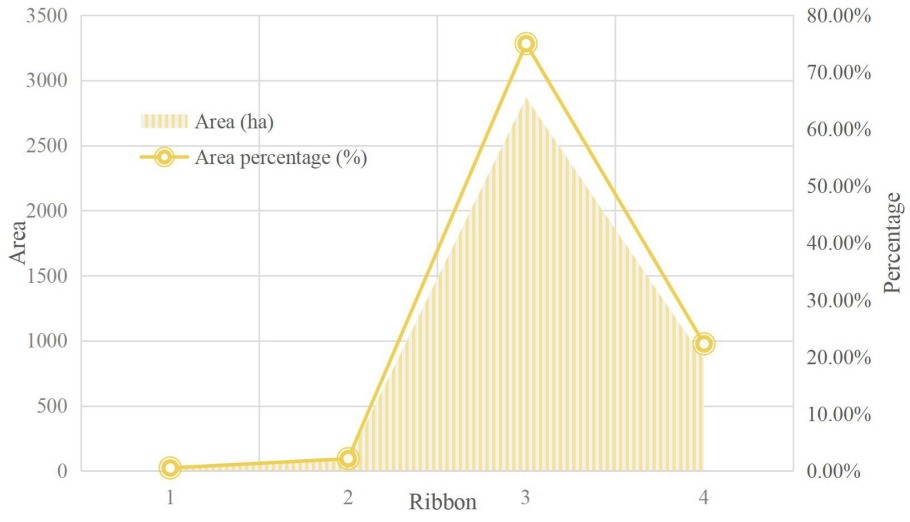


FIGURE 4
Analysis of functional zoning of forest park.

resource advantages of eco-tourism in the National Forest Park and realize the coordination and unity of various resource protection and planning and construction, the existing natural resources in the National Forest Park from water resources, climate resources, dynamic plant resources, etc. The comprehensive evaluation system of eco-tourism resources is systematically constructed from three directions. Combined with the specific conditions of national forest parks, it is further subdivided into two-level evaluation indicators such as biological resource quality, environmental resource quality, and regional and location conditions. The Delphi method, the expert evaluation method, is mainly used, and the research foundation and practical experience accumulated by experts in ecotourism planning are used to judge the grades of ecotourism resources at all levels, and then the national forest parks are classified according to the evaluation standards of individual tourism resources. The ecotourism landscape resources are classified (see Figure 3).

Experiment analysis

Comprehensive ecological characteristic analysis

The division of ecological function zones is mainly based on the types of the ecosystems of suburban forest parks and the differences in ecological functions. In Figure 4, according to the analysis of the functional area division of the forest park, in order to realize the coordination of elimination protection and planning and construction, it is necessary to make full use of the impact of Ecotourism of the National Forest Park. In Table 1, differences in ecological functions are shown. For suburban forest parks, functional zoning can control tourists' tour routes, and more importantly, protect ecologically sensitive areas, which play a key role in the sustainable development of regional ecosystems and the comprehensive improvement of the ecological environment.

Perception differences of different habitat types

In Table 2, it can be seen that the structure of forest recreation space plays a leading role in improving air quality, and coniferous forests have a more significant improvement in

TABLE 1 Comprehensive ecological characteristics of different ecological sensitivities.

Ecological grade	1	2	3
Category	Basic suitable land	Suitable land	Very suitable
Land use direction	Restricted planning	Proper planning	Unrestricted planning
Ecological sensitivity	Moderately sensitive	Mildly sensitive	insensitive
Comprehensive ecological characteristics	The terrain is low, the undulations are small, and it is far away from the ecological conservation area.	Mostly agricultural and forestry land	The terrain is low and flat, with abundant tourist attractions and good infrastructure.

TABLE 2 Differences in nature perception and air quality perception in different habitat types.

	Habitat type	Natural perception	Air quality perception
1	Deciduous broad-leaved forest	1.179	2.295
2	Coniferous and broadleaf mixed forest	1.285	2.349
3	Artificial coniferous forest	1.357	3.286
4	Primeval coniferous pine forest	1.697	2.152
5	Other coniferous forests	1.267	2.600

TABLE 3 Functional zoning table of forest park.

Serial number	Ribbon	Area (ha)	The main function	Area percentage (%)
1	Manage service area	20.11	Reception service, customer distribution, administrative management	0.53%
2	Core landscape area	82.73	Ecological tourism, cultural tourism, religious pilgrimage	2.16%
3	General recreation area	2867.24	Ecotourism, forest health, religious pilgrimage	75.00%
4	Ecological conservation area	852.92	Water conservation, soil, and water conservation, ecological maintenance	22.31%

The total planned area of the forest park is 3,823 ha.

air quality than broad-leaved forests. Samples of lower habitat quality confirmed these patterns, but users in higher quality habitats were more aware of the natural habitat itself.

The results show that aesthetic perception, emotion and natural perception are important factors affecting the cognitive mode of habitat space; the perception of soundscape is driven by the wet habitat; the recognition of the specific parts of Korean pine promotes the user's perception of the educational function of the habitat; Historic landscape sketches will evoke public perception of historical heritage.

Functional division of forest park

The details are shown in Table 3.

As can be seen from the above picture, as a strictly protected area, the ecological conservation area basically does not conduct scenic spot development and resource mining, nor is it open to tourists. The total area is 852.92 ha, accounting for 22.31% of the total area of the forest park. The Forest Park Ecological Conservation Area is strictly in accordance with the regulations, and there are no tourist attractions open to tourists, so that the ecology of the ecological conservation area can develop sustainably and healthily.

Conclusion

Eco-tourism design planning must be combined with management to improve the management level, clarify the development of forest parks and the short-term and long-term planning and design. Only good design combined with excellent management methods can maximize the benefits of ecotourism; reduce human interference to forest parks. For areas that have been damaged by humans, ecological restoration and reconstruction measures are taken to reconstruct the rich forest ecological landscape structure. At the same time, try to concentrate and narrow the scope of the construction area, pay attention to the protection of ecologically sensitive areas, and maintain the balance of the ecological system in the ecological conservation area; improve the risk awareness of tourists, and avoid the risk of sudden natural or man-made tourism disasters. Implement the strategy of ecological sustainable development of tourism, increase planning and publicity, attach importance to the shaping of tourism products with tourism image and characteristics of suburban forest parks, and expand the tourism market; it is necessary to mobilize the subjective initiative of residents in scenic spots to participate in tourism and promote local economic development, not only considering immediate interests, but also focus on long-term development.

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/s.

Author contributions

JZ was the experimental designer and the executive of the experimental research of this research. JS completed the data analysis and wrote the first draft of the manuscript. Both authors contributed to the article and approved the submitted version.

Conflict of interest

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

Publisher's note

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

References

Alkhalifa, I. S., and Almogren, A. S. (2020). NSSC: Novel Segment based Safety Message Broadcasting in Cluster based Vehicular Sensor Network. *IEEE Access* 8, 34299–34312 doi: 10.1109/ACCESS.2020.2974157

Awasthi, U., and Bollas, G. M. (2020). Sensor network design for smart manufacturing – Application on precision machining. *IFAC-PapersOnLine* 53, 11440–11445. doi: 10.1016/j.ifacol.2020.12.581

Baghaee, H. R., Mlaki, D., Nikolovski, S., and Dragicevic, T. (2020). Anti-Islanding Protection of PV-based Microgrids Consisting of PHEVs using SVMs. *IEEE Trans. Smart Grid* 11, 483–500. doi: 10.1109/TSG.2019.2924290

Blake, C., Christensen, M. N., and Kovács, K. T. (2021). Molecular Aspects of Plant Growth Promotion and Protection by Bacillus subtilis. *Mol. Plant Microbe Inter.* 34, 15–25. doi: 10.1094/MPMI-08-20-0225-CR

Khodadadiangostar, A., Rathnayake, T., Tennakoon, R. B., and Bab-Hadiashar, A. (2020). Centralized Cooperative Sensor Fusion for Dynamic Sensor Network With Limited Field-of-View via Labeled Multi-Bernoulli Filter. *IEEE Trans. Signal Processing* 69, 878–891. doi: 10.1109/TSP.2020.3048595

Manoj, M., Das, A. R., Chandran, A., and Mandal, S. (2020). Antecedents of environmental engagement and environmental learning behaviour. *J. Hosp. Tour. Insights* 3, 431–450. doi: 10.1108/JHTI-01-2020-0001

Martin, J. G., Maestre, J. M., and Camacho, E. F. (2021). Spatial irradiance estimation in a thermosolar power plant by a mobile robot sensor network. *Solar Energy* 220, 735–744. doi: 10.1016/j.solener.2021.03.038

Nobar, S. K., Mansourkiaie, F., and Ahmed, M. H. (2020). Packet Dropping Minimization in Energy Harvesting-Based Wireless Sensor Network With Linear Topology. *IEEE Access* 8, 38682–38691. doi: 10.1109/ACCESS.2020.2975489

Palani, U., Amuthavalli, G., and Alamelumangai, V. (2020). Secure and load balanced routing protocol in wireless sensor network for disaster management. *IET Inf. Secur.* 14, 513–520. doi: 10.1049/iet-ifs.2018.5057

Sambo, D. W., Forster, A., Yenke, B. O., Sarr, I., Gueye, B., and Dayang, P. (2020). Wireless Underground Sensor Networks Path Loss Model for Precision Agriculture (WUSN-PLM). *IEEE Sens. J.* 20, 5298–5313. doi: 10.1109/JSEN.2020.2968351

Venkataravanappa, V., Sonavane, P., Reddy, C., and Reddy Krishna, M. (2020). Associations of papaya leaf curl virus and betasatellite with yellow vein disease of pot marigold (*Calendula officinalis* L.). *Arch. Phytopathol. Plant Protection* 52, 1–9. doi: 10.1080/03235408.2019.1704350

Younus, M. U., Khan, M. K., and Anjum, M. R. (2020). Optimizing the Lifetime of Software Defined Wireless Sensor Network via Reinforcement Learning. *IEEE Access* 9, 259–272. doi: 10.1109/ACCESS.2020.3046693

Advantages of publishing in Frontiers

OPEN ACCESS
Articles are free to read for greatest visibility and readership

FAST PUBLICATION
Around 90 days from submission to decision

HIGH QUALITY PEER-REVIEW
Rigorous, collaborative, and constructive peer-review

TRANSPARENT PEER-REVIEW
Editors and reviewers acknowledged by name on published articles

Frontiers
Avenue du Tribunal-Fédéral 34
1005 Lausanne | Switzerland

Visit us: www.frontiersin.org
Contact us: frontiersin.org/about/contact

REPRODUCIBILITY OF RESEARCH
Support open data and methods to enhance research reproducibility

DIGITAL PUBLISHING
Articles designed for optimal readership across devices

FOLLOW US
@frontiersin

IMPACT METRICS
Advanced article metrics track visibility across digital media

EXTENSIVE PROMOTION
Marketing and promotion of impactful research

LOOP RESEARCH NETWORK
Our network increases your article's readership