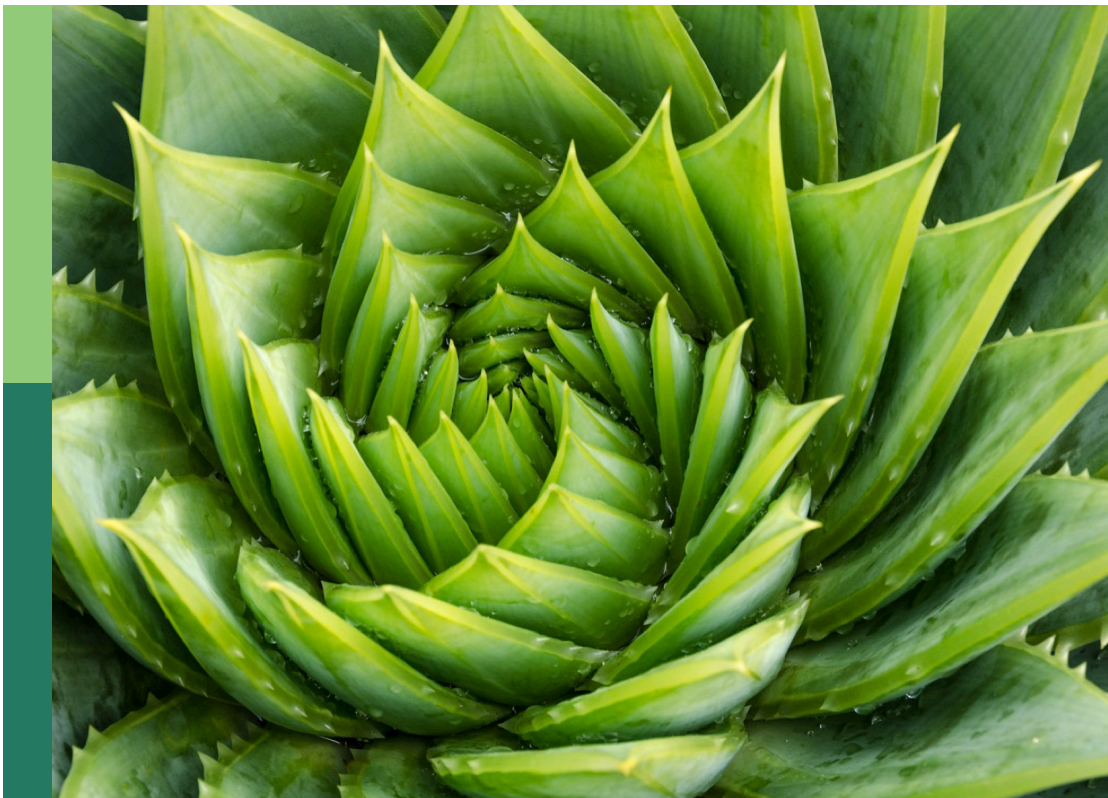
frontiers | Research Topics

Rapid and nondestructive detection techniques for agri-product quality

Edited by
Te Ma, Xi Tian, Qiang Lyu and Lie Deng

Published in
Frontiers in Plant Science



ISSN 1664-8714
ISBN 978-2-8325-5313-8
DOI 10.3389/978-2-8325-5313-8

About Frontiers

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

Frontiers journal series

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

Dedication to quality

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

What are Frontiers Research Topics?

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

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

Rapid and nondestructive detection techniques for agri-product quality

Topic editors

Te Ma — Nagoya University, Japan

Xi Tian — Intelligent Equipment Research Center, Beijing Academy of Agriculture and Forestry Sciences, China

Qiang Lyu — Southwest University, China

Lie Deng — Southwest University, China

Citation

Ma, T., Tian, X., Lyu, Q., Deng, L., eds. (2024). *Rapid and nondestructive detection techniques for agri-product quality*. Lausanne: Frontiers Media SA.

doi: 10.3389/978-2-8325-5313-8

Table of contents

frontiers | Frontiers in Plant Science

TYPE Review
PUBLISHED 16 August 2023
DOI 10.3389/fpls.2023.1240361

OPEN ACCESS

EDITED BY
Lie Deng,
Southwest University, China

REVIEWED BY
Yi Yang,
Beijing Academy of Agricultural and Forestry Sciences, China
Kusumiyati Kusumiyati,
Padjadjaran University, Indonesia

*CORRESPONDENCE
Byoung-Kwan Cho
chobk@cnu.ac.kr

RECEIVED 14 June 2023
ACCEPTED 27 July 2023
PUBLISHED 16 August 2023

CITATION
Aline U, Bhattacharya T, Faqeerzada MA, Kim MS, Baek I and Cho B-K (2023) Advancement of non-destructive spectral measurements for the quality of major tropical fruits and vegetables: a review. *Front. Plant Sci.* 14:1240361. doi: 10.3389/fpls.2023.1240361

Advancement of non-destructive spectral measurements for the quality of major tropical fruits and vegetables: a review

Umuhoza Aline[1], Tanima Bhattacharya[1], Mohammad Akbar Faqeerzada[2], Moon S. Kim[3], Insuck Baek[3] and Byoung-Kwan Cho[1,2]*

[1]Department of Agricultural Machinery Engineering, Chungnam National University, Daejeon, Republic of Korea, [2]Department of Smart Agricultural Systems, Chungnam National University, Daejeon, Republic of Korea, [3]Environmental Microbial and Food Safety Laboratory, Agricultural Research Service, United States Department of Agriculture, Beltsville, MD, United States

The quality of tropical fruits and vegetables and the expanding global interest in eating healthy foods have resulted in the continual development of reliable, quick, and cost-effective quality assurance methods. The present review discusses the advancement of non-destructive spectral measurements for evaluating the quality of major tropical fruits and vegetables. Fourier transform infrared (FTIR), Near-infrared (NIR), Raman spectroscopy, and hyperspectral imaging (HSI) were used to monitor the external and internal parameters of papaya, pineapple, avocado, mango, and banana. The ability of HSI to detect both spectral and spatial dimensions proved its efficiency in measuring external qualities such as grading 516 bananas, and defects in 10 mangoes and 10 avocados with 98.45%, 97.95%, and 99.9%, respectively. All of the techniques effectively assessed internal characteristics such as total soluble solids (TSS), soluble solid content (SSC), and moisture content (MC), with the exception of NIR, which was found to have limited penetration depth for fruits and vegetables with thick rinds or skins, including avocado, pineapple, and banana. The appropriate selection of NIR optical geometry and wavelength range can help to improve the prediction accuracy of these crops. The advancement of spectral measurements combined with machine learning and deep learning technologies have increased the efficiency of estimating the six maturity stages of papaya fruit, from the unripe to the overripe stages, with F1 scores of up to 0.90 by feature concatenation of data developed by HSI and visible light. The presented findings in the technological advancements of non-destructive spectral measurements offer promising quality assurance for tropical fruits and vegetables.

KEYWORDS

non-destructive measurement, spectral measurements, quality parameters, tropical fruits and vegetables, rapid measurement

1 Introduction

Tropical fruits and vegetables are agricultural crops that are typically grown in tropical regions where the climate is warm, with temperatures ranging from 20 to 35^{0}C (Bahadur et al., 2020). Tropical regions are found amidst the tropics of Cancer and Capricorn, and encompass equatorial zones in Oceania, Asia, Africa, Central and South America, and the Caribbean (Zakaria, 2023). Crops grown naturally in such weather conditions provide essential minerals, water, fiber, and vitamins that contribute significantly to the well-being of humans by safeguarding against ailments such as diabetes, hypertension, and cancer (Emelike and Akusu, 2019).

The agricultural revolution and the adaptation of numerous tropical plants to regions outside of their natural range have muddied their classification, and little is known about what properly defines and distinguishes tropical fruits and vegetables from their temperate counterparts (Indiarto, 2020). Fernandes et al. (Fernandes et al., 2011) described crop classification according to size, acidity, seed type, and bearing. Included among alkaline crops are apples, bananas, peaches, cherries, persimmon, and litchi (Fernandes et al., 2011). Acidic crops include strawberry, orange, kiwi, pineapple, lemon, star fruit, and logan, whereas sub-acidic examples are mango, pear, blackberry, papaya, blueberry, cherimoya, and mulberry (Fernandes et al., 2011). Chakraborty et al. (Chakraborty et al., 2014) agreed and structured the classification of tropical fruits based on that of Fernandes. Sarkar et al. (Sarkar et al., 2018) reported classification system according to maturity stage by means of ethylene gas emission and respiration rate, including both climacteric and non-climacteric tropical produce (Sarkar et al., 2018). Tropical climacteric produce such as avocado, apple, pear, mango, papaya, broccoli, banana, kiwi, and tomato undergoes maturation in correlation with an escalation in their respiration rate and the release of ethylene gas (Indiarto, 2020), whereas tropical non-climacteric crops such as grape, berry, citrus, litchi, strawberry, raspberry, pumpkin, watermelon, cucumber, and pineapple do not undergo an elevation in their respiration rate as they reach maturity (Indiarto, 2020). The contrasting report of Retamales et al. (Retamales, 2011) centers around the production of temperate crops worldwide. In this report, apple, raspberry, pear, peach, kiwi, blueberry, strawberry and plum were considered as temperate fruits (Retamales, 2011). In addition, Benichou et al. (Benichou et al., 2018) have also classified temperate fruits as tree (apple, plum, pear and peach), vine (grape and kiwi), and small fruits such as raspberry, blueberry and currant (Benichou et al., 2018).

Papaya, pineapple, avocado, mango, and banana are considered to be major tropical fruits globally (Mukhametzyanov et al., 2022). According to a market review prediction for the years 2013 to 2022 by the Food and Agriculture Organization of the United Nations (FAO), the most exported tropical fruits globally from Central America and the Caribbean, South America and Asia, Africa, and others in millions of tons were papaya, pineapple, avocado and mango with 3.7, 3.2, 2.3, and 2.1, respectively (Altendorf, 2019). On the other hand, recent data have shown that global vegetable production increased by 68% between 2000 and 2021 (FAO, 2022). Because of the continuous and emergent demand for tropical fruits and vegetables worldwide, the present emphasis is on quality assurance in relation to end-user inclinations and commercial standards (Silva and Abud, 2017). The quality of tropical fruits and vegetables is characterized by both external and internal parameters (Jha and Matsuoka, 2000). External parameters namely color, defects, size and shape depend on not only the appearance of the product, but also on the standards set (Cubero et al., 2016), whereas internal parameters such as nutritional value, internal defects, flavor, and texture are subjective to physicochemical composition and climate change (Zainalabidin et al., 2019). The quality of fruits and vegetables influences consumer preference and is directly or indirectly linked with further value-addition and processing technologies (James et al., 2010).

Several studies have identified postharvest losses as the most prominent factor among the origins of crop quality deterioration (Porat et al., 2018; Etana, 2019; Ahmad et al., 2021). Adding to that, high temperature and relative humidity are mentioned in the biological and chemical degradation of produce freshness, which affects sweetness, flavor, weight, turgor, and nutritional value (Elik et al., 2019). However, past reports indicated that low-temperature cooling systems and edible coating materials can be used to maintain and monitor the quality of these crops (Mendy et al., 2019; Jodhani and Nataraj, 2021). Conventional methods relying on the quantification of different quality traits such as dry matter content, oil content, and moisture content have also been reported in the study of quality parameters of fruits and vegetables; however, these methods were found to be undesirable, destructive, time-consuming, and labor-intensive (Magwaza and Tesfay, 2015; Kyriacou and Rouphael, 2018). Therefore, the application of non-destructive bio-sensing methods as a promising alternative for evaluating the value of tropical produce has been adopted (Ndlovu et al., 2022; Okere et al., 2022).

Computer vision and popular pre-trained convolutional neural network (CNN) models have been used as recognition systems to sort and grade different fruits and vegetables, especially in supermarkets, regarding their variety and species (Dubey and Jalal, 2012). However, computer vision can only assess external quality attributes due to the lack of spectral information (Rahman and Cho, 2016; Bhargava and Bansal, 2021). Acoustic emission technology involves the mechanical destruction of produce when subjected to mechanical or thermal stimulus (Aboonajmi et al., 2015) and is not appropriate for all categories of fruits and vegetables (Adedeji et al, 2020). Extensive works have been published on the evaluation of fruits and vegetables by spectral measurements such as Fourier transform infrared (FTIR) spectroscopy (Egidio et al., 2009), Near-infrared (NIR), Raman spectroscopy (Pandiselvam et al., 2022), and hyperspectral imaging (HSI) (Wang and Zhai, 2018). Generally, these reports have concentrated on the utilization of spectral measurements for determining targeted quality parameters of a particular fruit or vegetable variety. For instance, visible and near-infrared spectroscopy was used to investigate the internal browning in mango fruits (Gabriëls et al., 2020). Ali et al. (Ali et al., 2023) investigated FTIR, NIR, and machine vision in the quality

monitoring of pineapples. Metlenkin et al. (Metlenkin et al., 2022) distinguished Hass avocado fruits by defects using hyperspectral imaging (HSI). The question revolves around the practical utilization of these approaches and the challenges associated with improving data processing speed and in-line implementation (Cortés et al., 2019; Si et al., 2022). Quick hardware and software are required to fulfill the demands of swift analysis for extensive hyperspectral datasets (Xu et al., 2023) and machine learning algorithms, especially those relying on deep learning act as black boxes rather than using interpretability models for high-stakes decisions (Caceres-Hernandez et al., 2023).

The present review highlights the current advances in non-destructive spectral measurements for quality assessment, specifically for major tropical fruits and vegetables. The quality parameters of these tropical produces are covered first. The discussion on each of the spectral measurements, the tropical crops used, and the specific findings obtained from various studies, which are summarized in Table 1, follows and can deliver valuable information on the capabilities and efficiency of these techniques. In addition, the merits and demerits of each of these spectral measurements, which are presented in Table 2, will guide future researchers in selecting the proper evaluation method when evaluating the quality of tropical produces. To facilitate comprehension and quick understanding of key terminologies involved, the list of abbreviations and definitions contained in the paper is presented in Table 3.

2 Quality inspection of Tropical fruits and vegetables

Quality inspection is the process of evaluating specific parameters of fruits and vegetables to ensure required quality standards (Phey et al., 2020). The intention of quality inspection is to detect any internal or external characteristics that can aid in identifying both standard quality parameters and defects or non-conformities that can affect the safety of fruits and vegetables or

TABLE 1 A comparison of the application of various non-destructive spectral measurements in the quality assessment of tropical fruits and vegetables.

Measurement	Tropical produce	Parameter	Data analysis	Performance (Accuracy)	Reference
FTIR, FTNIR	Pineapple	SSC TA PH	PCA	SD=0.17 SD=0.11 SD=0.13	(Egidio et al., 2009)
Vis–NIR, ML	Mango	Color	PLS, ANN	80%	(Gabriëls et al., 2020)
HSI	Avocado	Defects	PCA, PLS-DA, SIMCA	99.9%	(Metlenkin et al., 2022)
NIR	Mango	Firmness	PCA,MPLS	R^2 = 0.88 R^2 = 0.85	(Flores et al., 2008)
NIR	Papaya	Starch SSC	PLS	R=0.90 R=0.90	(Purwanto et al., 2015)
Vis–NIR	Pineapple	Nitrates	PLSR	R=0.95	(Srivichien et al., 2015)
HSI	Potato	SSC	PLSR	R^2p=0.963	(Su and Sun, 2019)
FTIR	Banana	Maturity	PLS	R^2 = 0.83	(Zhang et al., 2021)
ATR-FTIR, ML	Banana	Ripening	PCA	96.0%	(Sinanoglou et al., 2023)
NIR	Avocado	Moisture content Dry matter	PLS	RPD= 2.00 RPD=2.13	(Olarewaju et al., 2016)
NIR	Mango	Maturity	MLR, PLS	Rc=0.74 Rv=0.68	(Jha et al., 2014)
NIR	Banana	TSS PH	PLS	R^2 = 0.81 R^2 = 0.69	(Ali et al., 2018)
NIR, HSI	Sweet potatoes	Variety identification	PLSDA	R^2 = 0.893	(Su et al., 2019)
NIR	Mango	Firmness	iPLSR	R^2c = 0.75 R^2p = 0.75	(Mishra et al., 2020)
Raman	Cassava	Starch adulteration	OC-SVM/SIMCA	86.9%	(Cardoso and Jesus Poppi, 2021)
Vis–NIR	Pineapple	Nitrate	PLSR	R= 0.95	(Srivichien et al., 2015)

(Continued)

TABLE 1 Continued

Measurement	Tropical produce	Parameter	Data analysis	Performance (Accuracy)	Reference
HSI	Banana	SSC TA	PLS/iPLS/PLSDA	R^2 = 0.64 R^2 = 0.59	(Chu et al., 2022)
NIR–HSI	Pineapple	Water activity	PLSR	Rp= 0.72	(Aozora et al., 2022)
HSI, ML, DL	Papaya	Maturity	DCNN	F1 = 0.91	(Garillos-Manliguez and Chiang, 2021)
Raman	Sweet potato	Moisture and carotenoids	PLSR&PCA	R^2 = 0.90(hot air) R^2 = 0.88 (microwave)	(Sebben et al., 2018)
Raman	Potato	Grading	PLSDA	≈100%	(Morey et al., 2020)
HSI	potato	Bruises	SVMM	87.88%	(Ye et al., 2018)
SWIR–HSI	Potato	Black spot	PLSDA	98.56%	(López-Maestresalas et al., 2016)
Raman	Mango	Carotenoids	–	R= 0.9618	(Bicanic et al., 2010)
Vis-NIR-HSI	Avocado	Nutrients (Fatty acids)	PLSR	R^2 = 0.79(flesh) R^2 = 0.62(skin)	(Kämper et al., 2020)
NIR–HSI	Mango	Defects	K-NN	97.95%	(Rivera et al., 2014)
HSI	Banana	Grading	CNN/MLP	98.45%	(Mesa and Chiang, 2021)

their usability in particular functions such as diets, trade, and industrial chains (Kirezieva et al., 2013).

2.1 External quality of tropical fruits and vegetables

The appearance of fruits and vegetables is a sensory attribute that directly influences the perceived worth of the produce for consumers (Zhang et al., 2014). The external quality of tropical crops is indicated by a number of factors, including size, shape, color, and external defects, as shown in Table 4 (Ganiron, 2014). The size and shape are two complementary factors that differ depending on the variety of the plant and are both assessed in relation to market grading standards (Abbaszadeh et al., 2013). The size is determined by measuring area, perimeter, length, and width, which is more complex due to the morphological irregularities of tropical crops natural state (Cubero et al., 2011). Moreda et al. (Moreda et al., 2009) described some non-invasive systems for assessing the size of fruits and vegetables. The systems are based on (1) measuring the volume of the gap between the fruit and the outer casing of an embracing gauge; (2) measuring the distance between a radiation source and the fruit contour, where this distance is computed from the time of flight (TOF) of the propagated waves; (3) light obstruction by barriers or blockades of light; (4) 2D and 3D machine vision systems (Moreda et al., 2009).

Wang et al. (Wang et al., 2017) evaluated mango size by RGB–D (depth) imaging and time-of-flight camera imaging system. The camera-to-fruit distance was determined using three methods for fruit sizing from images: stereo vision camera, RGB–D camera and a time-of-flight laser rangefinder (Wang et al., 2017). The obtained length and width values were good with RMSE of 4.9mm and 4.3mm respectively. It is cost-effective and simple to use; however, it pertains non-occluded fruit only and cannot be utilized in direct sunlight (Wang et al., 2017). Neupane et al. (Neupane et al., 2022) replicated the work of Wang by suggesting the use of partly occluded fruit. To obtain the linear length of the fruits, bounding box dimensions of an instance segmentation model (Mask R-CNN) was applied to canopy images (Neupane et al., 2022). The findings were good with RMSE values of 4.7 mm and 5.1 mm for Honey Gold and Keitt mango varieties, respectively (Neupane et al., 2022). Sanchez et al. (Sanchez et al., 2020) investigated spectroscopic and depth imaging techniques combined with machine vision to estimate the length, width, thickness, and volume of sweet potato and potato. When the correct size group was graded, the method had a high accuracy of 90% (Sanchez et al., 2020).

Color is an external quality trait that depends on the maturity of produce and is subjective to internal features such as taste, perception, and pleasantness of fruits and vegetables (Yahaya et al, 2017). Calorimeters evaluate color by measuring the typical surface area of the product and detects the color space values L*, a*, and b* which are based on the human color perception theory (Aguilar-Hernández et al., 2021). The capability of infrared thermal imaging approaches was investigated in the measurement of pineapple color. In this investigation, the L*, a*, and b* mean values for calorimeter increased by ($P < 0.05$) (Ali et al., 2022). The optical fiber sensors mounted with RGB LEDs were also used to evaluate the color of mangoes, giving R^2 = 0.879 (Yahaya et al., 2011).

External defects include the evidence of rot, bruising, crushing, shriveling, and wilting due to water loss which impact market value and the price of the fruits and vegetables (Raj and Suji, 2019). These defects can be recognized and monitored through the appearance of the crop by qualified personnel relying on subjective evaluation,

TABLE 2 Merits and demerits of non-destructive spectral measurements in the quality control of tropical fruits and vegetables.

Technique	Merits	Demerits	References
FTIR	No sample preparation.	Single beam and double beam for scattering device.	(Lan et al., 2020)
	Fast and easy to perform.	Difficulty in obtaining representative background.	
	Capability to measure many parameters at the same time.	Hard to read the interferogram if the Fourier transform is not performed first to generate the spectrum.	
	Good signal-to-noise ratio		
	Suitability for both quantitative and qualitative analyses.		
NIR	Real-time analysis.	Limited penetration depth.	(Srivichien et al., 2015), (Arendse et al., 2021)
	Can evaluate multiple components concurrently.	Time-consuming calibration procedure.	
	Fast acquisition of spectra.	Complex signal interpretation	
	Minimal sample preparation required.		
Raman	Vibrational and complementary.	Weak Raman scattering.	(Wang et al., 2021), (Li et al., 2016)
	Fast, Simple, sensitive, and selective technique.	Fluorescence interference.	
	Capability to monitor water-rich molecules.	Low reproducibility.	
	High spatial resolution.	Redundant data set. Costly Raman system.	
	Detects the spatial distribution of the molecules.	Relatively low operational speed	
HSI	Detect both spectral and spatial details.	Costly and complex data.	(Chandrasekaran et al., 2019), (Rajkumar et al., 2012)
	Concurrent assessment of many parameters.	Advanced hardware and software required.	
	Available in different algorithms.	Requires chemometrics techniques to extract relevant information.	

TABLE 3 List of abbreviations and acronyms used in the paper.

Abbreviation	Definition	Abbreviation	Definition
FTIR	Fourier transform infrared	CNN	Convolutional Neural Network
NIR	Near-infrared	TOF	Time of flight
HSI	Hyperspectral imaging	TSS	Total soluble solids
SSC	Soluble solid content	RGB–D imaging	Red, Green, Blue–Depth imaging
ASC	Added sugar content	PLS	Partial least squares
^{0}C	Degrees Celsius	RMSE	Root mean square error
FAO	Food and Agriculture Organization	YOLO	You Only Look Once
R-CNN	Regions with convolutional neural networks	ATR	Attenuated total reflectance
L*, a*, and b*.	Lightness, redness or greenness, and yellowness	MLR	Multivariate linear regression
LED	Light-emitting diode	IR	Infrared region
R^2	Determination coefficient	iPLSR	Interval partial least squares regression
TA	Total acidity	OC-SVM	One-class support vector machine
Vis–NIR	Visible–near-infrared spectroscopy	SIMCA	Soft independent modelling by class analogy

(Continued)

TABLE 3 Continued

Abbreviation	Definition	Abbreviation	Definition
R	Coefficient of correlation	SERS	Surface-Enhanced Raman Spectroscopy
PLSR	Partial least squares regression	RMSEP	Root mean square error of prediction
R^2P	Correlation of prediction	Rp	Coefficient of prediction
MIR	Mid-infrared	DT	Decision trees
FIR	Far-infrared	RNN	Recurrent neural network
ANN	Artificial neural network	PLSDA	Partial least square discriminant analysis
GA	Genetic algorithm	VGG	Visual Geometry Group
FL	Fuzzy logic	ResNet	Deep Residual Learning for Image Recognition
ANFIS	Adaptive neuro-fuzzy inference system	ResNeXt	Aggregated Residual Transformations for Deep Neural Networks
ML	Machine learning	DCNN	Deep convolutional neural network
DL	Deep learning	RPD	Residual predictive deviation
LDA	Linear discriminant analysis	F1 scores	Performance of Precision and recall
SVM	Support vector machine	MLP	Multilayer Perception
K-NN	K-nearest neighbors	PCA:	Principal component analysis
ELM	Extreme learning machine	MPLS:	Modified partial least square
RMSEC	Root mean square error of calibration	SD:	Standard deviation
Rc	Correlation coefficient for calibration	Rv	Correlation coefficient for validation

which may result in human errors (Ali et al., 2023). Sahu et al. (Sahu and Potdar, 2017) proposed a digital image analysis algorithm for detecting exterior defects in mango fruit. Surface defects such as scars and black patches were used to detect defective mango fruits, and were recognized by extracting the contours of damaged areas (Sahu and Potdar, 2017). The damaged area was then filled to identify its location in the image as the basis for discrimination. Sahu and colleagues achieved good accuracy but advocated the use of optimal and adaptive threshold approaches for segmenting mango fruits from image backgrounds (Sahu and Potdar, 2017).

2.2 Internal quality of tropical fruits and vegetables

The internal qualities of fruits and vegetables are also termed hidden qualities and are determined by texture, nutrients, internal defects, and flavor, as presented in Table 5 (Shewfelt, 2014). Different fruits and vegetables usually have different textures, which are characterized by their firmness, crispness, and crunchiness (Fillion and Kilcast, 2002). The assessment of fruit and vegetable firmness, a vital quality characteristic related to texture, can be achieved through sensory measurements (Magwaza and Opara, 2015). The texture is measured with a penetrometer by putting a probe tip installed on the texture analyzer into fruit tissue at a specific speed and depth so as to exert the most force (Ali et al., 2017). Uarrota et al. (Uarrota and Pedreschi, 2022) used a non-destructive texture analyzer to determine the firmness of avocado under different storage conditions. Enough data were required to construct the best model allowing an extension to the model firmness of avocado (Uarrota and Pedreschi, 2022). Kasim et al. (Kasim et al., 2021) compared laboratory-based (305-1713 nm) and portable-based (740-1070 nm) NIR spectrometers to determine mango firmness (Kasim et al., 2021). The results showed that portable and laboratory-based NIR instruments performed similar in respect of R^2p. Compared to the laboratory-based instrument, the RMSEP of the portable NIR was higher (Kasim et al., 2021).

Nutritional value, such as the sugar content related with vitamins and minerals, comprises the main constituents of soluble

TABLE 4 The external quality parameters of tropical fruits and vegetables.

External quality	Indicators	References
Size	Area, perimeter, length, and width	(Cubero et al., 2011), (Sanchez et al., 2020)
Shape	Mass, volume, spherical coefficient, density, and geometric mean diameter	(Cubero et al., 2011), (Golmohammadi and Afkari-Sayyah, 2013)
Color	Maturity, uniformity, and intensity	(Yahaya et al, 2017), (Ali et al., 2022)
External defects	Bruising, crushing, shriveling, and wilting	(Ali et al., 2023), (Raj and Suji, 2019)

TABLE 5 The internal quality parameters of tropical fruits and vegetables.

Internal quality	Indicator	References
Texture	Firmness, crispness, and juiciness	(Fillion and Kilcast, 2002), (Magwaza and Opara, 2015)
Nutrients	Chemical compositions (vitamins, sugars, proteins, and functional properties)	(Leiva-Valenzuela et al., 2013), (Aziz et al., 2021)
Internal defect	Internal cavity, water core, and rot	(Yahaya et al, 2017), (Ruiz-Altisent et al., 2010)
Flavor	Sweetness, sourness, saltiness, and bitterness	(Yahaya et al, 2017), (Zhu et al., 2020)

solids content (SSC), total soluble solids (TSS), and total acidity (TA) (Leiva-Valenzuela et al., 2013). Aziz et al. (Aziz et al., 2021) evaluated the relationship between TSS and the capacitance of papaya using capacitance-sensing techniques (Aziz et al., 2021). A refractometer was used as part of a destructive technique to predict the reference values of moisture and TSS content. Capacitive sensing was then tested as non-destructive approach for the evaluation of output voltage and capacitance of papaya (Aziz et al., 2021). Aziz observed a good correlation between destructive and non-destructive techniques, with R^2 of 0.9434 and 0.9177 for moisture and TSS content, respectively (Aziz et al., 2021). The usefulness of NIR spectroscopy was demonstrated in the determination of starch and soluble solid contents of papaya (Purwanto et al., 2015). Srivichien and colleagues tested the nitrates in pineapples using Vis–NIR (600-1200 nm) spectroscopy, yielding an R value of 0.95 (Srivichien et al., 2015). However, due to the big size and the change in nitrate levels, many scans were needed on different areas of pineapple (Srivichien et al., 2015). In the study to predict starch content of sweet potatoes and potatoes, hyperspectral imaging was applied by Su et al. (Su and Sun, 2019). Su developed partial least squares regression (PLSR) models at full-wavelength referring to spectral profiles and observed reference values, resulting in a high accuracy and an R^2P of 0.963 (Su and Sun, 2019).

Internal defects are detected as internal injury such as rot and water core inside the flesh of the fruits and vegetables due to postharvest problems(Ruiz-Altisent et al., 2010). Flavor or taste is defined by the sugar (sweetness), acidity (sourness), bitterness, and saltiness perceived by the tongue and nose (Zhu et al., 2020). It is, therefore, measured subjectively through oral testing or smelling, or by the conventional technical quantification of compounds such as liquid and gas chromatography (Yahaya et al, 2017). Korean universities conducted research on the taste and odor properties of broccoli using electronic sensors (Hong et al., 2022). For electronic tongue analysis, thermal processing boosted sourness and umami tastes while decreasing saltiness, sweetness, and bitterness (Hong et al., 2022). Therefore, the capability of non-destructive spectral measurement methods to assess inside parameters is important to maintain the flesh quality of tropical fruits and vegetables.

3 Non-destructive spectral measurements for the quality evaluation of tropical fruits and vegetables

Non-destructive techniques for quality monitoring of tropical fruits and vegetables refer to the process of inspecting their external and internal properties without causing damage or changing their physical and internal status (El-Mesery et al., 2019). The potential for employing spectral measurement approaches in the quality control of fruits and vegetables is growing enormously (Escárate et al., 2022). The reason is that these approaches are non-destructive, fast and accurate, capable for both quantitative and qualitative analysis, thereby requiring minimal sample preparation (Cozzolino, 2022). We divided non-destructive spectral measurements into two categories: (1) spectral-based approaches (FTIR, NIR, and Raman spectroscopy) and (2) imaging-based approaches (HSI), as shown in Figure 1.

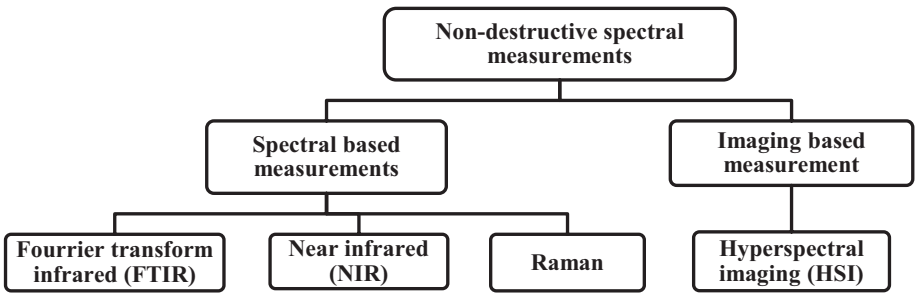


FIGURE 1
The schematic diagram of commonly used non-destructive spectral measurements.

3.1 Spectral-based approaches

Spectral measurement refers to effective techniques used to study the quality parameters of various agricultural materials including tropical fruits and vegetables by investigating light, sound, or particles that are emitted, absorbed, or scattered during measurement (Pathare and Rahman, 2022). Spectroscopic techniques based on FTIR, NIR, and Raman have been successful and popular in the detection of quality parameters of fruits and vegetables (Dasenaki and Thomaidis, 2019). Various research works have used spectral techniques focusing on fruits and vegetables, such as in the fast determination of the sugar and acid composition of citrus (Clark, 2016), assessment of primary sugars and amino acids in raw potato tubers (Ayvaz et al., 2015), and determination of nutrients and moisture content of fruits and vegetables (Sirisomboon, 2018). Quality parameters of tropical crops can be assessed by one of—or a sequence of—the above complementary techniques, which are distinguished depending on the infrared region (IR) they occupy and the molecular vibrations they detect (Bureau et al., 2019). The infrared region of the electromagnetic spectrum, presented in Figure 2, is separated into three sections, namely near-infrared (NIR), mid-infrared (MIR), and far-infrared (FIR) (Yeap and Hirasawa, 2019). Mango maturity has been predicted using the near-infrared (NIR) spectral region of 1200-2200 nm (Jha et al., 2014). The mid-infrared (MIR) spectral range of from 2500 to 25000 nm has been used in the prediction of banana maturity and geographical origin by Zhang et al. (Zhang et al., 2021), and in the measurement of soluble solids, total acids, and total anthocyanin in berries (Clark et al., 2018). Far-infrared (FIR) ranges have often been reported to be between 25000 and 300000 nm (Larkin, 2017). However, FIR applications are not clearly defined and are limited due to challenges in developing FIR instrumentation; furthermore, the band assignments of low-frequency vibrational modes are not straightforward (Ozaki, 2021). These spectral ranges are based on their relationship to the visible spectrum, which falls between 380 and 780 nm (Su and Sun, 2018).

3.1.1 Fourier transform infrared spectroscopy

FTIR is a form of vibrational spectroscopy that uses light interference to identify the chemical composition of scanned samples by producing infrared absorption or emission spectra (Larkin, 2017). On the electromagnetic spectrum, FTIR operates in the MIR region (2500 to 25000nm) and generates fruit or vegetable chemical profile by capturing the principle vibrational and rotational stretching modes of molecules (Lohumi et al., 2015). FTIR spectroscopy comprises of an infrared light source, interferometer, sample, and detector, shown in Figure 3. The principal part is the interferometer which is made up of three components: the beam splitter, collimator, and the two mirror (fixed and movable mirror) (Patrizi and Cumis, 2019). When the radiation from the light source passes through the collimator, strikes the beam splitter which ideally divide it into two beams. The first beam hits the static mirror, and is reflected back; while the second hits the movable mirror where it enters through the sample toward the detector (Blum and Harald, 2012).

The FTIR associated with attenuated total reflection (ATR-FTIR) has recently gained importance (Chan and Kazarian, 2016). The ATR works under the principle of total internal reflectance where infrared light interacts with the sample of high refractive index only at the point where infrared light is reflected (Ryu et al., 2021). Unlike transmission methods, the ATR-FTIR technique can be used to study solid, liquid, and paste samples with minimal sample preparation (Glassford et al., 2013).The combination of ATR-FTIR and chemometrics was promising in the assessment of added sugar content, (ASC), total soluble solids (TSS) and real juice content (RJC) of fresh and commercial mango juice (Jha and Gunasekaran, 2010). PLS and MLR models resulted into accuracy of 0.99 and 0.98 respectively (Jha and Gunasekaran, 2010). Canteri et al. (Canteri et al., 2019) have used ATR-FTIR to evaluate the cell wall compositions of 29 species of fruits and vegetables as freeze-dried powders and alcohol-insoluble solids. The results were accurate, with determination coefficient $R^2 \geq 0.9$ (Canteri et al., 2019). Recently, Sinanoglou et al. (Sinanoglou et al., 2023) conducted the evaluation of both peel and fresh banana ripening stage by ATR-FTIR, along with image analysis, discriminant and statistical analysis (Sinanoglou et al., 2023). The computed features were accurate enough to separate ripening stages; however, monitoring of the banana ripening process was highly reliant on the instrument employed for image analysis such as digital cameras, smartphones, and electronic noses (Sinanoglou et al., 2023).

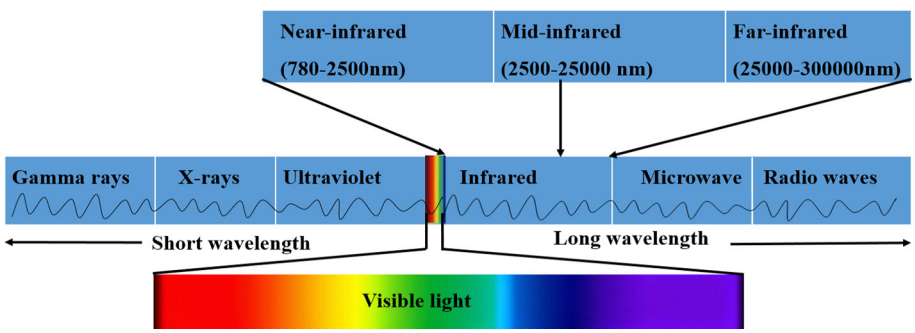


FIGURE 2
Modified diagram showing the infrared regions of the electromagnetic spectrum (Yeap and Hirasawa, 2019), (Aboud et al., 2019).

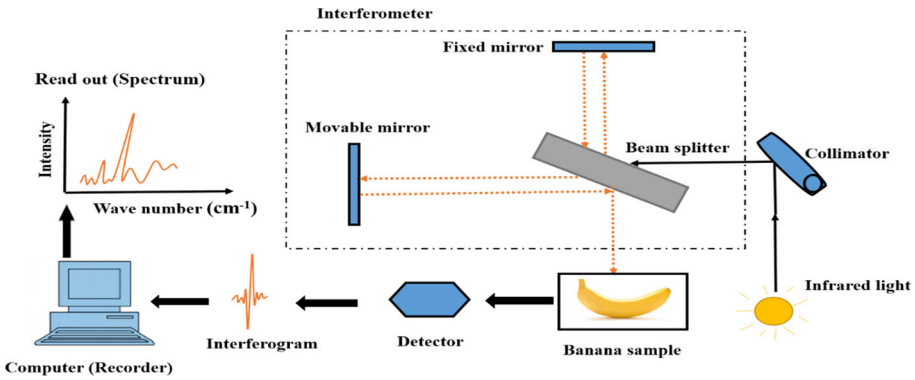


FIGURE 3
Modified diagram of FTIR spectroscopy taking banana as sample (Patrizi and Cumis, 2019).

3.1.2 Near-Infrared spectroscopy

NIR is used to rapidly ascertain the chemical constitution of materials according to overtones and harmonic or combination bands of specific functional groups (Kusumaningrum et al., 2018). Those overtones and combinations of vibrational bands characterized by C–H, O–H, and N–H are gained by NIR in the wavelength region of 780-2500nm (Ozaki et al., 2006). Tsuchikawa et al. (Tsuchikawa et al., 2022) described NIR as a spectroscopic method that is suitable for samples of high water content, including fruits and vegetables (Tsuchikawa et al., 2022). NIR spectroscopy consists of a light source, sample accessory, monochromator (grating), detector, and optical components such as lenses and optical fibers, as shown in Figure 4 (Lee et al., 2011).

The illumination of NIR light to the sample occurs in three ways: reflectance, interactance and transmittance (Wang et al., 2014). According to Hong and colleagues, reflectance employs high light energy, has no contact with the fruit surface, and the source and sensor are placed at a specified angle (Hong and Chia, 2021). Specular reflectance and diffuse reflectance are two types of reflectance measurement. Specular reflectance, which occurs when the incident and reflected angles are same, detects nothing from the inside part of the fruit (Hong and Chia, 2021); While the capacity of diffuse reflectance to constrain light dispersion into solid samples allows the acquisition of interior fruit information (Tang et al., 2022). Mango TSS, firmness, TA, and ripeness index (RPI) were effectively measured by NIR diffuse reflectance, with R^2 of 0.9; 0.82; 0.74; and 0.8, respectively. The effect of changes in physicochemical properties of mango during ripening, on the other hand was highlighted (Rungpichayapichet et al., 2016). Kusumiyati et al. (Kusumiyati and Suhandy, 2021) also evaluated TSS and Vitamin C using the same fruit and NIR spectra acquisition mode. The diffuse reflectance spectra were documented and found to be in relation with TSS, vitamin C (Kusumiyati and Suhandy, 2021).

Delwiche et al. (Delwiche et al., 2008) demonstrated the use of near infrared interactance (750-1088nm) to determine mango ripeness, SSC and other sugars. The mango sample was placed in contact with the probe in which the top of mango upwardly points the probe. The R^2 was 0.77; 0.75; 0.67; and 0.70 for SSC, sucrose, glucose, and fructose, respectively. Sugars such as sucrose indicates mango sweetness, fructose and glucose increases during ripening while acidity decreases (Delwiche et al., 2008). Transmission mode in which the light source and sensor are opposite to each other, employs low light intensity to reflect the inner parameters and is performed with no contact on the fruit (Nicolaï et al., 2007).

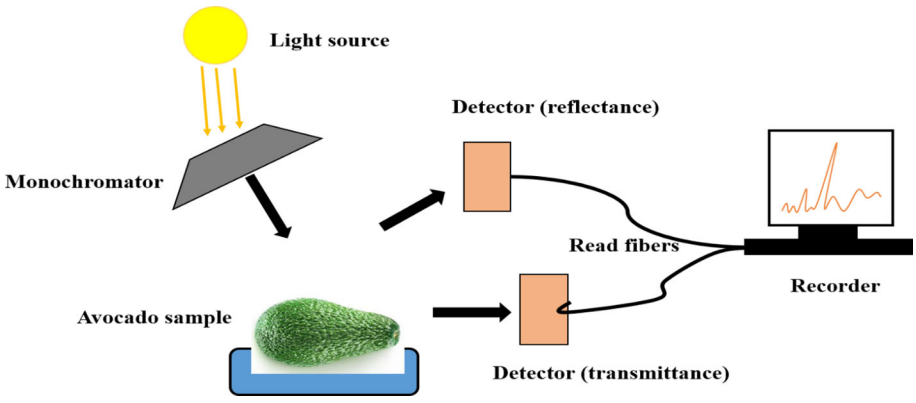


FIGURE 4
Modified diagram of NIR spectroscopy, taking avocado as sample (Chandrasekaran et al., 2019).

Transmission might be done partially or fully. Though, the difference between partial transmission and diffuse reflectance remains undetermined since both evaluate the radiation that partly enters the sample and diffusely reproduced to the sensor (Hong and Chia, 2021). The fruit with large seed such as mango was reported to be hard to measure in the full transmission due the low signal to noise ratio (Greensill and Walsh, 2000). Subedi at al. (Subedi and Walsh, 2011) detected the TSS and DM of mesocarp tissue of banana and mango by partial transmittance. Mango DM gave R^2cv =0.75 while banana performance negatively influenced by the thickness of the peel. The TSS results on mango was good in ripe and poor in ripening stage with R^2cv > 0.75 and R^2p < 0.75 respectively. The results were consistent with those of Rungpichayapichet et al. (Rungpichayapichet et al., 2016) and were found to be caused by the physiological factors of Mango, banana, and other tropical fruits which can change their starch content as they ripe (Subedi and Walsh, 2011).

Several studies have highlighted the potentials of NIR spectroscopy to monitor the internal and external characteristics of tropical fruits and vegetables, including the following: maturity prediction of avocado and mango (Olarewaju et al., 2016; S. N. Jha et al., 2014), total soluble solids and pH of banana (Ali et al., 2018), and variety identification in sweet potatoes (Su et al., 2019). However, the irregular thick skin of pineapple and chemical complexity of large seeded mango was the main difficulty to Guthrie et al. (Guthrie and Walsh, 1997) in the measurement of SSC by NIR reflectance (760-2500nm). The penetration depth of NIR light into a thick-rind avocado 38 mm in diameter and 10 mm in thickness was investigated for the maturity evaluation of avocado using an NIR spectrometer (800–2400 nm) (Olarewaju et al., 2016). The models for estimating oil content, were acceptable, however were not accurate, with an RPD value of less than 1.0 and an R^2 value of 0.58 (Olarewaju et al., 2016). Arendse et al. (Arendse et al., 2018) informed the limited accuracy of NIR for internal quality assessment of fruits and vegetables with thick rinds such as banana, avocado and pineapple due to inadequate penetration depth (Arendse et al., 2018). Therefore, future studies can consider the appropriate selection of NIR optical geometry and wavelength range to improve the prediction accuracy of thick rind tropical crops (Pratiwi et al., 2023).

NIR spectral data inevitably holds overlay information of numerous organic compounds at global wavelengths, making the use of global spectroscopic regions problematic rather than specific wave bands (Lin and Yibin, 2009). Therefore, a combination of algorithms and chemometrics with NIR spectroscopy is now being used to meet this demand, balance data redundancy and complexity, and collect spectral information (Guan et al., 2019; Yang et al., 2021). Portable NIR spectroscopy was used to assess mango firmness during ripening (400–1130 nm) (Mishra et al., 2020). Pre-processing was done Savitzky–Golay filter, and iPLSR model was found to provide better predictive modeling, with an R^2p of 0.75 and an RMSEC of 5.92 $Hz^2g^{2/3}$ compared to the standard PLSR model, which had an R^2p of 0.67 and an RMSEC of 6.88 $Hz^2g^{2/3}$. For the firmness in mango fruit, spectral intervals 743-770 nm and 870-905 nm were found to be the accurate predictors (Mishra et al., 2020).

3.1.3 Raman spectroscopy

Raman is another form of vibrational spectroscopy that uses laser beams to interact with materials and operates in the infrared region of the electromagnetic spectrum from 2500 to 25000 nm (Siesler et al., 2008). Though Raman and MIR spectroscopy methods use high levels of energy to detect molecular vibrations, Raman spectroscopy excels at equal vibrations of nonpolar sets, while MIR spectroscopy excels at the unequal vibrations of polar sets (Campanella et al., 2021). Raman spectroscopy consists of a monochromatic laser, wavelength separator, and a detector, as presented in Figure 5 (Qin et al., 2019). When the laser beam illuminates the sample, the photons that constitute the light are absorbed, transmitted, or scattered by the sample in different directions before reaching the detector (Larkin, 2017). Absorption and transmission are linked with the infrared spectra (IR), while scattering is associated with the Raman spectra (Jones et al., 2019). Rostron et al. (Rostron et al., 2016) defined scattered photons in two different ways namely Rayleigh (elastic) scattering and Raman (inelastic) scattering (Larkin, 2017). Rayleigh (elastic) scattering

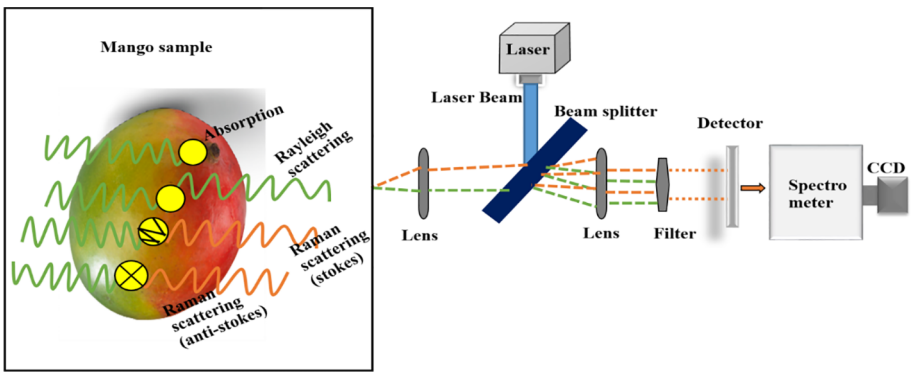


FIGURE 5
Modified diagram of Raman spectroscopy, taking mango as sample (Lohumi et al., 2015).

occurs when the photons scattered are equal to those illuminated to the sample; while Raman (inelastic) scattering is due to the transfer of energy between photons and the sample under testing (Lu, 2017).

Raman spectroscopy is suitable for investigating carotenoids in various plants, including carrots (Lawaetz et al., 2016), tomatoes (Hara et al., 2018), plant cells (Baranska et al., 2011), and mango (Bicanic et al., 2010). Furthermore, Raman has been applied as a clean and fast approach to assess cassava starch adulteration (Cardoso and Jesus Poppi, 2021). Two chemometrics models, namely one-class support vector machines (OC-SVMs) and soft independent modelling by class analogy (SIMCA), were used and compared statistically. The OC-SVM results outperform those of SIMCA, with an accuracy of 86.9% (Cardoso and Jesus Poppi, 2021). Surface-enhanced Raman spectroscopy (SERS) was used as a method that applies Raman spectroscopy in conjunction with nanotechnology for the fast analysis of pesticide residues in mango (Pham et al., 2022). SERS results were good indicating that the residues in mango sample were in the suitable range (Pham et al., 2022). Morey et al. (Morey et al., 2020) used spatially offset Raman spectroscopy for potato varieties quality categorization and prediction of tuber cultivation source. This approach is fast since it can be used directly after potato harvesting (Morey et al., 2020).

3.2 Imaging-based approaches

Spectral imaging techniques are among the most effective detection methods because of their potential to obtain both spectral and spatial dimensions of produce simultaneously during measurement (Liu et al., 2017). Regarding spatial dimensions, external attributes such as size, shape, appearance, and color can be evaluated, while with spectral analysis, internal features such as chemical composition can be measured (Pu et al., 2015). A number of imaging techniques use two-dimensional geometry according to the fusion and luminance of color maps (Lu et al., 2014), while others involve the use of three-dimensional sensors such as RGB and hyperspectral images (Barnea et al., 2016) to provide a high fruit and vegetable recognition accuracy (Nyarko et al., 2018).

3.2.1 Hyperspectral imaging techniques

In agriculture and food systems, hyperspectral imaging is a powerful system that joins two aspects of imaging and spectroscopy to attain a three-dimensional (3D) hypercube data form and analyzes a broad spectrum at each pixel instead of assigning only main RGB colors (red, green, and blue) (Khan et al., 2021). The hypercube consists of 3D images characterized by 2D spatial and 1D spectral dimension or wavelength (Tang et al., 2022). Hyperspectral imaging employs more than ten contiguous wavelengths or narrow bands in which each pixel has a full continuous spectrum (Elmasry et al., 2019). To take sample images, the hyperspectral imaging set up can be in the reflectance, transmittance, and interactance which differs in their lighting configuration during crops measurements (Pan et al., 2017). The reflectance geometry is appropriate for assessing the external quality of products, whereas the transmittance performs better in measuring the internal components in relatively translucent membranes (Li et al., 2018). The HSI system comprises of four main components: (1) an imaging unit, (2) illumination (light source), (3) a sample stage, and (4) a computer, as presented in Figure 6 (Pu et al., 2015). The light source is divided into illumination and excitation sources for spectral imaging applications. Broadband lights are commonly used as an illumination source for reflectance and transmittance, whereas narrowband lights are for the excitation source (Qin et al., 2013). The lighting devices produce light that illuminates the sample. The camera transports chemical information as well as light from the light source. The wavelength dispersion device, which can be a grating or a prism, divides the light into different wavelengths and directs the dispersed light to the sensor (Wu and Sun, 2013). Aozora et al. (Aozora et al., 2022) studied the efficiency of hyperspectral imaging (935–1720 nm) in the evaluation of water activity in dehydrated pineapple. The accuracy of the tested model showed good accuracy, with 0.72 and 0.0054 for Rp of and RMSEP respectively (Aozora et al., 2022).

3.2.1.1 Hyperspectral imaging Image generation modes

HSI generates image in three ways: whisk broom (point scanner), push broom (line scanner), and tunable filter (area

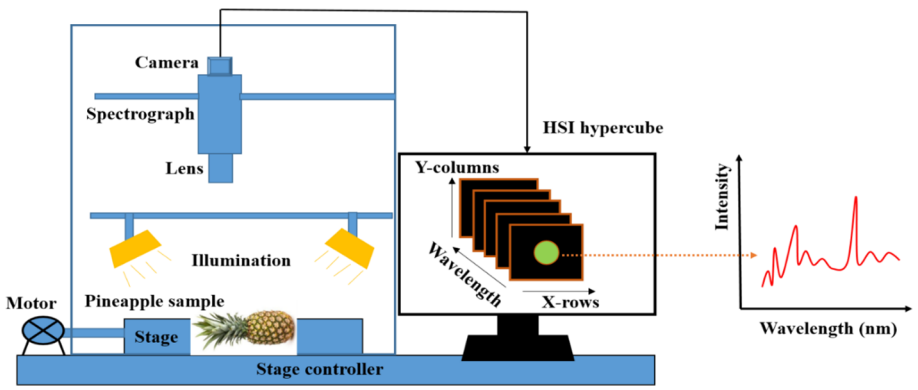


FIGURE 6
Modified diagram of Hyperspectral imaging, taking pineapple as sample (Li et al., 2018).

scanner) (ElMasry and Sun, 2010). The point scan excites only a single spot on the object's surface and the single pixel is recorded. The spectrum is taken at both positions by moving the sample symmetrically in two spatial dimensions, in order to get the full HSI image (Qin, 2012). However, to obtain good results this technique involves double scanning of the sample and hardware relocation which takes a lot of time to complete the measurement (Qin, 2012). The line scanner excites a line on the object and records the whole line of an image using a 2D dispersing element and 2D detector array. The object is moved line by line and the whole set of spatial–spectral data is gained. This approach has a higher acquisition rate but lower sectioning ability (Qin, 2010). The area scan employs spectral scanning techniques to stimulate the broad area on the surface of the fruit or vegetable, which is held fixed and a scan with full spatial information is achieved consecutively across the entire spectral range. This method is appropriate for applications where sample mobility is not necessary (Lu et al., 2017).

The hyperspectral imaging together with chemometrics models is an appealing option for dealing with large sets of complex, high-dimensional data (Lorente et al., 2012). Chu et al. (Chu et al., 2022) confirmed the efficacy of the HSI reflectance (386-1016 nm) wavelength region in combination with variable selection algorithms and chemometrics for predicting green banana maturity level and characterization of banana quality during maturation (Chu et al., 2022). The line scanning approach was adopted and the calibration models used were partial least squares (PLS) and interval PLS methods (Chu et al., 2022). These models obtained acceptable values R^2 = 0.64 and 0.59 for SSC and TA, respectively, whereas the models for chlorophyll and ΔE* were suitable only for sample screening with R^2 = 0.34 and 0.30, respectively (Chu et al., 2022). Chu reported the inclusion of more samples and different cultivars of banana for model improvement (Chu et al., 2022). Kämper et al. (Kämper et al., 2020) used Vis–NIR–HSI to measure nutrients in avocado fruit. PLSR was used to obtain the ratio of unsaturated to saturated fatty acids in avocado fruit with (R^2 = 0.79, RPD = 2.06) and (R^2 = 0.62, RPD = 1.48) for flesh images and skin images respectively (Kämper et al., 2020). The robust models for flesh images were R^2 = 0.67; 0.61; and 0.53, of oleic-to-linoleic acid ratio, boron (B) and calcium concentration (Ca) respectively, while for skin images was R^2 = 0.60 of boron (Kämper et al., 2020).

4 Advancement in non-destructive spectral measurements for tropical fruit and vegetable quality assessment

The rapid advancement of technology in the agricultural field has resulted in the combination of artificial intelligence with non-destructive spectral measurements for fruits and vegetables quality measurement (Hasanzadeh et al., 2022). Artificial intelligence models such as artificial neural networks (ANNs), genetic algorithms (GAs), fuzzy logic (FL), and adaptive neuro-fuzzy inference system (ANFIS) can assess multiple characteristics simultaneously (Homayoonfal et al., 2022). Salehi reviewed development of models used in the determination of fruits and vegetables quality (Salehi, 2020). ANNs, GAs, FL, and ANFIS detected defects, moisture content, and chilling injury of oranges, cherries, pomegranates, apples, peaches, avocados, button mushrooms, tomatoes, and potatoes (Salehi, 2020). Despite the fact that these models are typically constrained by normality, linearity, homogeneity, and variable independence, the ANFIS model outperforms others and can be successfully used in relevant research (Salehi, 2020).

Machine learning (ML) is a branch of artificial intelligence and an integral part of the development of many sensing technologies that are responsible for information retrieval, signal processing, and data analysis (Li et al., 2021). In recent decades, traditional algorithms such as linear discriminant analysis (LDA), support vector machines (SVMs), K-nearest neighbors (K-NN), naïve Bayes, extreme learning machines (ELMs), decision trees (DTs), and K-means clustering have been deployed (Fadchar and Dela Cruz, 2020). For instance, Rivera et al. (Rivera et al., 2014) used NIR–HSI and machine learning for the early detection of mechanical damage in mango. LDA, K-NN, naïve Bayes, ELMs, and DTs were used for categorization. Bayes failed, however (K-NN, ELM, DT, and LDA Title altered) results was more than 90%. The highest performance, achieved by K-NN, was 97.9% (Rivera et al., 2014).

The evolution of deep learning (DL) as a breakthrough machine learning method has been trending since 2017 due to the manual feature extraction of traditional machine learning methods (Yang and Xu, 2021) and limited performance of chemometrics models, such as spectral variability caused by sample and spectrometer heterogeneity, changing environmental conditions, and infrared spectral data with high noise, which hinder feature extraction using chemometrics models (Zhang et al., 2021). Deep learning is a subset of machine learning that use many neural network layers to extract complex feature representations with numerous levels of abstraction (Lecun et al., 2015). According to Kamilaris et al. (Kamilaris and Prenafeta-Boldú, 2018), convolutional neural network (CNN) and recurrent neural network (RNN) have been implemented for crop-type classification, counting produces, and locating their placement in the image using bounding boxes (Kamilaris and Prenafeta-Boldú, 2018). However, the RNN was found to perform better than the CNN because it considers not only space but also the time which helps to capture the time dimension (Kamilaris and Prenafeta-Boldú, 2018). Deep learning and machine learning technology-based spectral analysis has been used in the classification of three types of fruits (apple, lemon, and mango) by type of damage, type of goods, and whether the sample is raw in market, supermarket, wholesaler, and retailer applications (Bobde et al., 2021).

Garillos-Manliguez et al. (Garillos-Manliguez and Chiang, 2021) estimated six maturity stages of papaya fruit, from the unripe stage to the overripe stage, by feature concatenation of data obtained from visible light and HSI imaging (Garillos-Manliguez and Chiang, 2021). AlexNet, VGG16, VGG19, ResNet50, ResNeXt50, MobileNet, and MobileNetV2 architectures was then modified to apply multimodal data cubes made of RGB and hyperspectral data (Garillos-Manliguez and Chiang, 2021). Regarding classification of the six stages, these multimodal

variations can reach F1 scores of up to 0.90 and a 1.45% top-2 error rate. However, due to the small size of the images and the great depth of the CNNs, resulting in highly tightly tuned training variables, overfitting may arise. On the other hand, increasing image size results in insufficient memory faults (Garillos-Manliguez and Chiang, 2021).

Banana fruit was graded by Mesa et al. (Mesa and Chiang, 2021) using multi-input deep learning model with RGB and HSI. These models were able to categorize tier-based bananas by 98.45% and an F1 score of 0.97 with only few samples (Mesa and Chiang, 2021). However, this technique is expensive and time consuming due to the use of two cameras. The next studies instead, should consider the use of more improved camera systems with features that can extract both RGB and HSI simultaneously (Mesa and Chiang, 2021). Another study by Ucat and Cruz explored the use of image processing with a deep learning to grade banana according to their specifications (Ucat and Dela Cruz, 2019). The trained, validated, and test data by CNN model was more than 90% in all four classes of bananas (). The suggested CNN grading system in the tensor flow model can be commercially developed (Ucat and Dela Cruz, 2019).

Portable spectrometers and real-time online detection devices have recently developed for fruits and vegetables quality assessment. Portable devices are handheld, light weight, compact size and they are applied for in-field measurements (Sohaib et al., 2020). The combination of portable NIR device with MSC-PCA+LDA model was used to evaluate pineapple quality. These models were recommended to be developed in mobile phone while PLS regression model provided 85% accuracy (Amuah et al., 2019). Subedi et al. (Subedi and Walsh, 2020) evaluated three hand held portable near infrared spectroscopy (F750, Micro NIR and Scio v1.2) in the detection of dry matter content (DMC) in avocado fruit. The second derivative spectra were recorded for the intact and skin removed avocado fruit for reflectance and interactance optical geometry. The best results of prediction obtained from the F750 instrument using the interactance mode at 720-975 nm with R^2p of 0.71 and 0.88 for intact and skin removed fruits respectively (Subedi and Walsh, 2020). Real time monitoring device was designed as sensor which can function in all post-harvesting states to control the shelf life of fruits and vegetables such as lettuce. The device found to be the feasible for controlling the behavior of the crop during the post handling chain (Torres-Sánchez et al., 2020). Fruits and vegetables including banana, orange and apple were well sorted according to their external appearance by using real time online system with artificial intelligence (Tata et al., 2022). For quality categorization, machine learning models such as CNN and image processing were performed. This real time system was created in android and can be deployed in market robots where checking of huge number of products is required (Tata et al., 2022).

5 Conclusion and future prospects

Non-destructive spectral measurement has emerged as a prominent solution in the agricultural sector. With the introduction of spectral measurements, there has been rapid progress in analyzing both the internal and external characteristics of tropical fruits and vegetables in a low-cost, accurate, real-time, and fast manner (Ali et al., 2017). Techniques based on FTIR, NIR, and Raman spectroscopy require simple steps to prepare samples prior to analysis (Abbas et al., 2020). In contrast to other imaging techniques such as computer vision, acoustic approaches, electric noses, and fluorescence, HSI uses spectral and spatial data to assess different parameters concurrently (Lu et al., 2020). The spectral measurements presented in this review have shown potential applications for a diverse range of tropical fruits and vegetables for the monitoring and detection of quality attributes such as SSC, TSS, TA, color, size, defects, and texture, which is particularly important for fruit and vegetable processors, food safety agencies, and consumer demands.

Significant advancements in non-destructive spectral measurement technology have occurred recently, including the development of portable spectrometers for real-time and field applications. The combination of spectral measurements and chemometric techniques is a powerful tool for multivariate data analysis, mainly in the improvement of models needed for classification and estimation of quality. A practical case study of Metlenkin et al. (Metlenkin et al., 2022) in the identification and classification of Hass avocado defects before and after storage by HSI and chemometrics. The PLSDA and SIMCA were selected as chemometric methods for multivariate data discrimination and classification. To increase the final model accuracy the calibration was performed by selecting the region of interest. The results revealed the high potential of SIMCA during both modelling and test validation with 100% accuracy. Furthermore, the integration of spectral measurements with deep learning and machine learning technology is rapidly expanding in order to improve quality control accuracy while overcoming the challenges associated with chemometrics such as spectral variability, spectrometer heterogeneity, changing environmental conditions, and infrared spectral data with high noise. The revolution in agriculture and the adaptation of numerous tropical plants to regions outside of their natural range have muddied their classification, and little is known about what properly defines and distinguishes tropical fruits and vegetables from their temperate counterparts. Therefore, there is confusion associated with those studies that reported the classification of tropical fruits and vegetables as an important factor to consider when examining the distinctive quality indicators of these crops. Taking into accounts all of the merits and demerits of non-destructive spectral measurements for the quality monitoring of tropical fruits and vegetables, the use of an

adequate number of samples, different cultivars of the fruit and increasing the quality attributes to predict can help to develop robust models that emphasize the variability of tropical fruits and vegetables in terms of size and shape, skin thickness, and growing conditions.

Author contributions

Conceptualization: UA, B-KC. Methodology: UA, TB, MF, MK and IB. Investigation: UA, TB and B-KC. Writing and reviewing: UA, TB, MF and B-KC. Supervision: B-KC. All authors contributed to the article and approved the submitted version.

Funding

This work was supported by Korea Institute of Planning and Evaluation for Technology in Food, Agriculture, Forestry (IPET) through Smart Agri Products Flow Storage Technology Development Program, funded by Ministry of Agriculture, Food and Rural Affairs (MAFRA) (322051-05).

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

Abbas, O., Pissard, A., and Baeten, V. (2020). "Near-infrared, mid-infrared, and raman spectroscopy," in *Chemical analysis of food* (Amstardam: Elsevier), 77–134.

Abbaszadeh, R., Rajabipour, A., Ahmadi, H., Mahjoob, M. J., and Delshad, M. (2013). Prediction of watermelon quality based on vibration spectrum. *Postharvest Biol. Technol.* 86, 291–293. doi: 10.1016/j.postharvbio.2013.07.013

Aboonajmi, M., Jahangiri, M., and Hassan-Beygi, a. S. R. (2015). A review on application of acoustic analysis in quality evaluation of agro-food products. *J. Food Process. Preservation* 39 (6), 3175–3885. doi: 10.1111/jfpp.12444

Aboud, S. A., Altemimi, A. B., Al-hiiphy, A. R.S., Yi-chen, L., and Cacciola, F. (2019). A comprehensive review on infrared heating. *Molecules* 2, 1–20.

Adedeji, A. A., Ekramirad, N., Rady, A., Hamidisepehr, A., Donohue, K. D., Villanueva, R. T., et al(2020)Non-destructive technologies for detecting insect infestation in fruits and vegetables under postharvest conditions: A critical review*Foods* 9 (7), 1–285doi: 10.3390/foods9070927

Aguilar-Hernández, M. G., Núñez-Gómez, Dámaris, Forner-Giner, MaríaÁngeles, Hernández, F., Pastor-Pérez, JoaquínJ., and Legua, P. (2021). Quality parameters of spanish lemons with commercial interest. *Foods* 10 (1), 1–135. doi: 10.3390/foods10010062

Ahmad, K., Afridi, M., Khan, N. A., and Sarwar, A. (2021). Quality deterioration of postharvest fruits and vegetables in developing country Pakistan: A mini overview. *Asian J. Agric. Food Sci.* 9 (2), 83–90. doi: 10.24203/ajafs.v9i2.6615

Ali, M. M., Hashim, N., Aziz, S. A., and Lasekan, O. (2022). Quality prediction of different pineapple (Ananas comosus) varieties during storage using infrared thermal imaging technique. *Food Control* 138, 1–9. doi: 10.1016/j.foodcont.2022.108988

Ali, M. M., Hashim, N., Bejo, S. K., Jahari, M., and Shahabudin, N. A. (2023). Innovative non-destructive technologies for quality monitoring of pineapples: recent advances and applications. *Trends Food Sci. Technol.* 133, 176–188. doi: 10.1016/j.tifs.2023.02.005

Ali, M. M., Hashim, N., Bejo, S. K., and Shamsudin, R. (2017). Rapid and nondestructive techniques for internal and external quality evaluation of watermelons: A review. *Scientia Hortic.* 225, 689–699. doi: 10.1016/j.scienta.2017.08.012

Ali, M. M., Janius, R. B., Nawi, N. M., and Hashim, N. (2018). Prediction of total soluble solids and PH in banana using near infrared spectroscopy. *J. Eng. Sci. Technol.* 13 (1), 254–645.

Altendorf, (2019). Major tropical fruits. *Stat. Compendium Rome* 01, 18.

Amuah, C. L. Y., Teye, E., Lamptey, F. P., Nyandey, K., Opoku-Ansah, J., and Aduening., P. O. W. (2019). Feasibility study of the use of handheld NIR spectrometer for simultaneous authentication and quantification of quality parameters in intact pineapple fruits. *J. Spectrosc.*, 1–9. doi: 10.1155/2019/5975461

Aozora, D. W., Tantinantrakun, A., Thompson, A. K., and Teerachaichayut, S. (2022). Near infrared hyperspectral imaging for predicting water activity of dehydrated pineapples. *Res. Militaris* 12 (2022), 11–33. doi: 10.3390/foods12142793

Arendse, E., Fawole, O. A., Magwaza, L. S., and Opara., U. L. (2018). Non-Destructive prediction of internal and external quality attributes of fruit with thick rind: A review. *J. Food Eng.* 217, 11–23. doi: 10.1016/j.jfoodeng.2017.08.009

Arendse, E., Nieuwoudt, H., Magwaza, L. S., Nturambirwe, J. F. I., Fawole, O. A., and Opara., U. L. (2021). Recent advancements on vibrational spectroscopic techniques for the detection of authenticity and adulteration in horticultural products with a specific focus on oils, juices and powders. *Food Bioprocess Technol.* 14 (1), 1–225. doi: 10.1007/s11947-020-02505-x

Ayvaz, H., Santos, A. M., Moyseenko, J., Kleinhenz, M., and Rodriguez-Saona, L. E. (2015). Application of a portable infrared instrument for simultaneous analysis of sugars, asparagine and glutamine levels in raw potato tubers. *Plant Foods Hum. Nutr.* 70 (2), 215–205. doi: 10.1007/s11130-015-0484-7

Aziz, N. A. A., Jusoh, M. Z., and Rosman, R. (2021). "Relationship of total soluble solid (TSS) and capacitance value of papaya fruit using capacitive sensing technique," in *ISCI 2021 - 2021 IEEE symposium on computers and informatics*, Kuala Lumpur, Malaysia, 51–57. doi: 10.1109/ISCI51925.2021.9633402

Bahadur, L., Singh, A. D., and Singh, S. K. (2020). A review on successful protected cultivation of banana (Musa). *Plant Arch.* 20, 1570–1573.

Baranska, M., Baranski, R., Grzebelus, E., and ROman, M. (2011). *In situ* detection of a single carotenoid crystal in a plant cell using raman microspectroscopy. *Vibrational Spectrosc.* 56 (2), 166–695doi: 10.1016/j.vibspec.2011.02.003

Barnea, E., Mairon, R., and Ohad, B.-S. (2016). Colour-Agnostic shape-Based 3D fruit detection for crop harvesting robots. *Biosyst. Eng.* 146, 57–70. doi: 10.1016/j.biosystemseng.2016.01.013

Benichou, M., Ayour, J., Sagar, M., Alahyane, A., Elateri, I., and Aitoubahou, A. (2018). Postharvest technologies for shelf life enhancement of temperate fruits. *Postharvest Biol. Technol. Temperate Fruits*, 77–100. doi: 10.1007/978-3-319-76843-4_4

Bhargava, A., and Bansal, A. (2021). Fruits and vegetables quality evaluation using computer vision: A review. *J. King Saud Univ. - Comput. Inf. Sci.* 33 (3), 243–575. doi: 10.1016/j.jksuci.2018.06.002

Bicanic, D., Dimitrovski, D., Luterotti, S., Twisk, C., Buijnsters, J. G., and Dóka, Ottó (2010). Estimating rapidly and precisely the concentration of beta carotene in mango homogenates by measuring the amplitude of optothermal signals, chromaticity indices and the intensities of raman peaks. *Food Chem.* 121 (3), 832–385. doi: 10.1016/j.foodchem.2009.12.093

Blum, M. M., and Harald, J. (2012). Historical perspective and modern applications of attenuated total reflectance - fourier transform infrared spectroscopy (ATR-FTIR). *Drug Testing Anal.* 4 (3–4), 298–302. doi: 10.1002/dta.374

Bobde, S., Jaiswal, S., Kulkarni, P., Patil, O., Khode, P., and Jha, R. (2021). "Fruit quality recognition using deep learning algorithm," in *2021 International Conference on Smart Generation Computing, Communication and Networking (SMART GENCON)*, Pune, India, 1–5. doi: 10.1109/SMARTGENCON51891.2021.9645793

Bureau, S., Cozzolino, D., and Clark, C. J. (2019). Contributions of fourier-transform mid infrared (FT-MIR) spectroscopy to the study of fruit and vegetables: A review. *Postharvest Biol. Technol.* 148, 1–14. doi: 10.1016/j.postharvbio.2018.10.003

Caceres-Hernandez, D., Gutierrez, R., Kung, K., Rodriguez, J., Lao, O., Contreras, K., et al. (2023). Recent advances in automatic feature detection and classification of fruits including with a special emphasis on watermelon (Citrillus lanatus): A review. *Neurocomputing* 526, 62–79. doi: 10.1016/j.neucom.2023.01.005

Campanella, B., Palleschi, V., and Legnaioli, S. (2021). Introduction to vibrational spectroscopies. *ChemTexts* 7 (1), 1–21. doi: 10.1007/s40828-020-00129-4

Canteri, M. H. G., Renard, C. M. G. C., Bourvellec, C. Le, and Bureau, S. (2019). ATR-FTIR spectroscopy to determine cell wall composition: application on a large diversity of fruits and vegetables. *Carbohydr. Polymers* 212, 186–196. doi: 10.1016/j.carbpol.2019.02.021

Cardoso, K. V. G., and Jesus Poppi, R. (2021). Cleaner and faster method to detect adulteration in cassava starch using raman spectroscopy and one-class support vector machine. *Food Control* 125, 107917. doi: 10.1016/j.foodcont.2021.107917

Chakraborty, K., Saha, J., Raychaudhuri, U., and Chakraborty, R. (2014). Tropical fruit wines: A mini review. *Natural Products* 7 (10), 219–285.

Chan, K. L., and Kazarian, S. G. (2016). Attenuated total reflection fourier-transform infrared (ATR-FTIR) imaging of tissues and live cells. *Chem. Soc. Rev.* 45 (7), 1850–1645. doi: 10.1039/C5CS00515A

Chandrasekaran, I., Panigrahi, S. S., Ravikanth, L., and Singh, C. B. (2019). Potential of near-Infrared (NIR) spectroscopy and hyperspectral imaging for quality and safety assessment of fruits: an overview. *Food Analytical Methods* 12 (11), 2438–2585. doi: 10.1007/s12161-019-01609-1

Chu, X., Miao, Pu, Zhang, K., Wei, H., Fu, H., Liu, H., et al. (2022). Green banana maturity classification and quality evaluation using hyperspectral imaging. *Agric. (Switzerland)* 12 (4), 1–185. doi: 10.3390/agriculture12040530

Clark, C. J. (2016). Fast determination by fourier-transform infrared spectroscopy of sugar-acid composition of citrus juices for determination of industry maturity standards. *New Z. J. Crop Hortic. Sci.* 44 (1), 69–82. doi: 10.1080/01140671.2015.1131725

Clark, C. J., Cooney, J. M., Hopkins, W. A., and Currie, A. (2018). Global mid-infrared prediction models facilitate simultaneous analysis of juice composition from berries of actinidia, ribes, rubus and vaccinium species. *Food Analytical Methods* 11 (11), 3147–3605. doi: 10.1007/s12161-018-1296-9

Cortés, V., Blasco, J., Aleixos, N., Cubero, S., and Talens, P. (2019). Monitoring strategies for quality control of agricultural products using visible and near-infrared spectroscopy: A review. *Trends Food Sci. Technol.* 85, 138–148. doi: 10.1016/j.tifs.2019.01.015

Cozzolino, D. (2022). An overview of the successful application of vibrational spectroscopy techniques to quantify nutraceuticals in fruits and plants. *Foods* 11 (3), 1–11. doi: 10.3390/foods11030315

Cubero, S., Aleixos, N., Moltó, E., Gómez-Sanchis, J., and Blasco, J. (2011). Advances in machine vision applications for automatic inspection and quality evaluation of fruits and vegetables. *Food Bioprocess Technol.* 4 (4), 487–5045. doi: 10.1007/s11947-010-0411-8

Cubero, S., Lee, W. S., Aleixos, N., Albert, F., and Blasco., J. (2016). Automated systems based on machine vision for inspecting citrus fruits from the field to postharvest—a review. *Food Bioprocess Technol.* 9 (10), 1623–1395. doi: 10.1007/s11947-016-1767-1

Dasenaki, M. E., and Thomaidis, N. S. (2019). *Quality and authenticity control of fruit juices-a review*. *Molecules* 24, 1–35. doi: 10.3390/molecules24061014

Delwiche, S. R., Mekwatanakarn, W., and Wang, C. Y. (2008). Soluble solids and simple sugars measurement in intact mango using near infrared spectroscopy. *HortTechnology* 18 (3), 410–165. doi: 10.21273/horttech.18.3.410

Dubey, S. R., and Jalal, XXXA. S. (2012). Robust approach for fruit and vegetable classification. *Proc. Eng.* 38, 3449–3453. doi: 10.1016/j.proeng.2012.06.398

Egidio, V. Di, Sinelli, N., Limbo, S., Torri, L., Franzetti, L., and Casiraghi, E. (2009). Evaluation of shelf-life of fresh-cut pineapple using FT-NIR and FT-IR spectroscopy. *Postharvest Biol. Technol.* 54 (2), 87–925. doi: 10.1016/j.postharvbio.2009.06.006

Elik, A., Yanik, D. K., Istanbullu, Y., Guzelsoy, N. A., Yavu, A., and Gogus., F. (2019). Strategies to reduce post-harvest losses for fruits and vegetables. *Int. J. Sci. Technological Res.* 5 (3), 29–395. doi: 10.7176/jstr/5-3-04

Elmasry, G., Mandour, N., Al-Rejaie, S., Belin, E., and Rousseau, D. (2019). Recent applications of multispectral imaging in seed phenotyping and quality monitoring—An overview. *Sensors (Switzerland)* 19 (5), 1–325. doi: 10.3390/s19051090

ElMasry, G., and Sun, D.-W. (2010). "Principles of hyperspectral imaging technology," in *Hyperspectral imaging for food quality analysis and control* (San Diego: Elsevier), 3–43.

El-Mesery, H. S., Mao, H., and Abomohra, A. El F. (2019). Applications of non-destructive technologies for agricultural and food products quality inspection. *Sensors (Switzerland)* 19 (4), 1–235. doi: 10.3390/s19040846

Emelike, N. J. T., and Akusu, O. M. (2019). Quality attributes of jams and marmalades produced from some selected tropical fruits. *J. Food Process Technol.* 10 (5), 790. doi: 10.4172/2157-7110.1000790

Escárate, P., Farias, G., Naranjo, P., and Zoffoli, J. P. (2022). Estimation of soluble solids for stone fruit varieties based on near-infrared spectra using machine learning techniques. *Sensors* 22 (16), 1–115. doi: 10.3390/s22166081

Etana, M. B. (2019). A detailed review on common causes of postharvest loss and quality deterioration of fruits and vegetables in Ethiopia. *J. Biology Agric. Healthcare* 9 (7), 48–52. doi: 10.7176/jbah/9-7-07

Fadchar, N. A., and Dela Cruz, J. C. (2020). "A non-destructive approach of young coconut maturity detection using acoustic vibration and neural network," in *Proceedings - 2020 16th IEEE International Colloquium on Signal Processing and Its Applications, CSPA*, Langkawi, Malaysia, 136–140. doi: 10.1109/CSPA48992.2020.9068723

FAO (2022). *Agricultural production statistics 2000–2021. Agricultural production statistics 2000–2021* (Rome: FAO). doi: 10.4060/cc3751en

Fernandes, F. A. N., Rodrigues, S., Law, C. L., and Mujumdar, A. S. (2011). Drying of exotic tropical fruits: A comprehensive review. *Food Bioprocess Technol.* 4 (2), 163–855. doi: 10.1007/s11947-010-0323-7

Fillion, L., and Kilcast, D. (2002). Consumer perception of crispness and crunchiness in fruits and vegetables. *Food Qual. Preference* 13 (1), 23–295. doi: 10.1016/S0950-3293(01)00053-2

Flores, K., Sanchez, M. T., Perez-Marin, D. C., Lopez, M. D., Guerrero, J. E., and Garrido-Varo, A. (2008). Prediction of total soluble solid content in intact and cut melons and watermelons using near infrared spectroscopy. *J. Near Infrared Spectrosc.* 16 (2), 91–98. doi: 10.1255/jnirs.771

Gabriëls, S. H. E. J., Mishra, P., Mensink, M. G. J., Spoelstra, P., and Woltering, E. J. (2020). Non-destructive measurement of internal browning in mangoes using visible and near-infrared spectroscopy supported by artificial neural network analysis. *Postharvest Biol. Technol.* 166, 111206. doi: 10.1016/j.postharvbio.2020.111206

Ganiron, T. U. (2014). Size properties of mangoes using image analysis. *Int. J. Bio-Science Bio-Technology* 6 (2), 31–42. doi: 10.14257/ijbsbt.2014.6.2.03

Garillos-Manliguez, C. A., and Chiang, J. Y. (2021). Multimodal deep learning and visible-light and hyperspectral imaging for fruit maturity estimation. *Sensors (Switzerland)* 21 (4), 1–185. doi: 10.3390/s21041288

Glassford, S. E., Byrne, B., and Kazarian, S. G. (2013). Recent applications of ATR FTIR spectroscopy and imaging to proteins. *Biochim. Biophys. Acta - Proteins Proteomics.* 1834 (12), 2849–2858. doi: 10.1016/j.bbapap.2013.07.015

Golmohammadi, A., and Afkari-Sayyah, A. H. (2013). Long-term storage effects on the physical properties of the potato. *Int. J. Food Properties* 16 (1), 104–135. doi: 10.1080/10942912.2010.529978

Greensill, C. V., and Walsh, K. B. (2000). Remote acceptance probe and illumination configuration for spectral assessment of internal attributes of intact fruit. *Measurement Sci. Technol.* 11 (12), 1674–1845. doi: 10.1088/0957-0233/11/12/304

Guan, X., Liu, J., Huang, K., Kuang, J., and Liu, D. (2019). Evaluation of moisture content in processed apple chips using NIRS and wavelength selection techniques. *Infrared Phys. Technol.* 98, 305–310. doi: 10.1016/j.infrared.2019.01.010

Guthrie, J., and Walsh, K. (1997). Non-invasive assessment of pineapple and mango fruit quality using near infra-red spectroscopy. *Aust. J. Exp. Agric.* 37 (2), 253–263. doi: 10.1071/EA96026

Hara, R., Ishigaki, M., Kitahama, Y., Ozaki, Y., and Genkawa, T. (2018). Excitation wavelength selection for quantitative analysis of carotenoids in tomatoes using raman spectroscopy. *Food Chem.* 258, 308–313. doi: 10.1016/j.foodchem.2018.03.089

Hasanzadeh, B., Abbaspour-Gilandeh, Y., Soltani-Nazarloo, A., Hernández-Hernández, M., Gallardo-Bernal, Iván, and Hernández-Hernández, JoséL. (2022). Non-destructive detection of fruit quality parameters using hyperspectral imaging, multiple regression analysis and artificial intelligence. *Horticulturae* 8 (7), 1–16. doi: 10.3390/horticulturae8070598

Homayoonfal, M., Malekjani, N., Baeghbali, V., Ansarifar, E., Hedayati, S., and Jafari, S. M. (2022). Optimization of spray drying process parameters for the food bioactive ingredients. *Crit. Rev. Food Sci. Nutr.*, 1–41. doi: 10.1080/10408398.2022.2156976

Hong,, and Chia, K. S. (2021). A review on recent near infrared spectroscopic measurement setups and their challenges. *Measurement: journal of the international measurement confederation*. 171, 108732. doi: 10.1016/j.measurement.2020.108732

Hong, S. J., Yoon, S., Lee, J., Jo, S. M., Jeong, H., Lee, Y., et al. (2022). A comprehensive study for taste and odor characteristics using electronic sensors in broccoli floret with different methods of thermal processing. *J. Food Process. Preservation* 46 (4), 1–125. doi: 10.1111/jfpp.16435

Indiarto, R. (2020). Post-harvest handling technologies of tropical fruits: A review. *Int. J. Emerging Trends Eng. Res.* 8 (7), 3951–3957. doi: 10.30534/ijeter/2020/165872020

James, J. B., Ngarmsak, T, and Rolle, R. S. (2010). *Processing of fresh-cut tropical fruits and vegetables: A technical guide*. RAP Publication 2010/16.

Jha, S. N., and Gunasekaran, S. (2010). Authentication of sweetness of mango juice using fourier transform infrared-attenuated total reflection spectroscopy. *J. Food Eng.* 101 (3), 337–342. doi: 10.1016/j.jfoodeng.2010.07.019

Jha, S. N., Narsaiah, K., Jaiswal, P., Bhardwaj, R., Gupta, M., Kumar, R., et al. (2014). Nondestructive prediction of maturity of mango using near infrared spectroscopy. *J. Food Eng.* 124, 152–157. doi: 10.1016/j.jfoodeng.2013.10.012

Jha, S. N., and Matsuoka, T. (2000). Non-Destructive techniques for quality evaluation of intact fruits and vegetables. *Food Sci. Technol. Res.* 6 (4), 248–515. doi: 10.3136/fstr.6.248

Jodhani, K. A., and Nataraj, M. (2021). Synergistic effect of aloe gel (Aloe vera L.) and lemon (Citrus limon L.) peel extract edible coating on shelf life and quality of banana (Musa spp.). *J. Food Measurement Characterization* 15 (3), 2318–2285. doi: 10.1007/s11694-021-00822-z

Jones, R. R., Hooper, D. C., Zhang, L., Wolverson, D., and Valev, V. K. (2019). Raman techniques: fundamentals and frontiers. *Nanoscale Res. Lett.* 14 (1), 1–34. doi: 10.1186/s11671-019-3039-2

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

Kämper, W., Trueman, S. J., Tahmasbian, I., and Hosseini Bai, S. (2020). Rapid determination of nutrient concentrations in hass avocado fruit by vis/NIR hyperspectral imaging of flesh or skin. *Remote Sens.* 12 (20), 1–195. doi: 10.3390/rs12203409

Kasim, N. F. M., Mishra, P., Schouten, R. E., Woltering, E. J., and Boer, M. P. (2021). Assessing firmness in mango comparing broadband and miniature spectrophotometers. *Infrared Phys. Technol.* 115, 103733. doi: 10.1016/j.infrared.2021.103733

Khan, M. H., Saleem, Z., Ahmad, M., Sohaib, A., Ayaz, H., Mazzara, M., et al. (2021). Hyperspectral imaging-based unsupervised adulterated red chili content transformation for classification: identification of red chili adulterants. *Neural Computing Appl.* 33 (21), 14507–14215. doi: 10.1007/s00521-021-06094-4

Kirezieva, K., Jacxsens, L., Uyttendaele, M., Martinus, A. J. S., Boekel, V., and Luning, P. A. (2013). Assessment of food safety management systems in the global fresh produce chain. *Food Res. Int.* 52 (1), 230–425. doi: 10.1016/j.foodres.2013.03.023

Kusumaningrum, D., Lee, H., Lohumi, S., Mo, C., Kim, M. S., and Kwan Cho, B. (2018). Non-destructive technique for determining the viability of soybean (Glycine max) seeds using FT-NIR spectroscopy. *J. Sci. Food Agric.* 98 (5), 1734–1425doi: 10.1002/jsfa.8646

Kusumiyati, A. A. M., and Suhandy, D. (2021). Fast, simultaneous and contactless assessment of intact mango fruit by means of near infrared spectroscopy. *AIMS Agric. Food* 6 (1), 172–845. doi: 10.3934/AGRFOOD.2021011

Kyriacou, M. C., and Rouphael, Y. (2018). Towards a new definition of quality for fresh fruits and vegetables. *Scientia Hortic.* 234, 463–469. doi: 10.1016/j.scienta.2017.09.046

Lan, W., Renard, C. M. G. C., Jaillais, B., Leca, A., and Bureau, S. (2020). Fresh, freeze-dried or cell wall samples: which is the most appropriate to determine chemical, structural and rheological variations during apple processing using ATR-FTIR spectroscopy? *Food Chem.* 330, 127357. doi: 10.1016/j.foodchem.2020.127357

Larkin, P. (2017). *Infrared and raman spectroscopy: principles and spectral interpretation* (Amsterdam: Elsevier).

Lawaetz, A. J., Christensen, S. M. U., Clausen, S. K., Jørnsgaard, B., Rasmussen, SørenK., Andersen, S. B., et al. (2016). Fast, cross cultivar determination of total carotenoids in intact carrot tissue by raman spectroscopy and partial least squares calibration. *Food Chem.* 204, 7–13. doi: 10.1016/j.foodchem.2016.02.107

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

Lee, J.-D., Shannon, J.G., and Choung, M.-G. (2011). Application of nondestructive measurement to improve soybean quality by near infrared reflectance spectroscopy. *Soybean Appl. Technol.* 16, 287–304. doi: 10.5772/15842

Leiva-Valenzuela, G. A., Lu, R., and Aguilera, JoséM. (2013). Prediction of firmness and soluble solids content of blueberries using hyperspectral reflectance imaging. *J. Food Eng.* 115 (1), 91–985. doi: 10.1016/j.jfoodeng.2012.10.001

Li, Y. C., Khan, F., Jan, S. R. U., Verma, S., Menon, V. G., Kavita,, et al. (2021). A comprehensive survey on machine learning-based big data analytics for ioT-enabled smart healthcare system. *Mobile Networks Appl.* 26 (1), 234–525. doi: 10.1007/s11036-020-01700-6

Li, J. L., Sun, Da W., and Cheng, J. Hu (2016). Recent advances in nondestructive analytical techniques for determining the total soluble solids in fruits: A review. *Compr. Rev. Food Sci. Food Saf.* 15 (5), 897–9115. doi: 10.1111/1541-4337.12217

Li, R. Li, Wang, M., Liu, Y., Zhang, B., and Zhou, J. (2018). Hyperspectral imaging and their applications in the nondestructive quality assessment of fruits and vegetables. *Hyperspectral Imaging Agriculture Food Environ.* 27–63. doi: 10.1007/intechopen.72250

Lin, H., and Yibin, Y. (2009). Theory and application of near infrared spectroscopy in assessment of fruit quality: A review. *Sens. Instrumentation Food Qual. Saf.* 3 (2), 130–415. doi: 10.1007/s11694-009-9079-z

Liu, Y., Pu, H., and Sun, Da W. (2017). Hyperspectral imaging technique for evaluating food quality and safety during various processes: A review of recent applications. *Trends Food Sci. Technol.* 69, 25–35. doi: 10.1016/j.tifs.2017.08.013

Lohumi, S., Lee, S., Lee, H., and Cho, B. K. (2015). A review of vibrational spectroscopic techniques for the detection of food authenticity and adulteration. *Trends Food Sci. Technol.* 46 (1), 85–985. doi: 10.1016/j.tifs.2015.08.003

López-Maestresalas, A., Keresztes, J. C., Goodarzi, M., Arazuri, S., Jarén, C., and Saeys, W. (2016). Non-destructive detection of blackspot in potatoes by vis-NIR and SWIR hyperspectral imaging. *Food Control* 70, 229–241. doi: 10.1016/j.foodcont.2016.06.001

Lorente, D., Aleixos, N., Gómez-Sanchis, J., Cubero, S., García-Navarrete, O. L., and Blasco, J. (2012). Recent advances and applications of hyperspectral imaging for fruit and vegetable quality assessment. *Food Bioprocess Technol.* 5 (4), 1121–1142. doi: 10.1007/s11947-011-0725-1

Lu, (2017). *Light scattering technology for food property, quality and safety assessment* (Boca Raton, USA: Crc Press), 1–43.

Lu, N. S., Hu, Y., and Fu, H. (2014). Detecting citrus fruits with highlight on tree based on fusion of multi-map. *Optik* 125 (8), 1903–1975. doi: 10.1016/j.ijleo.2013.04.135

Lu, Y., Huang, Y., and Lu, R. (2017). Innovative hyperspectral imaging-based techniques for quality evaluation of fruits and vegetables: A review. *Appl. Sci. (Switzerland)* 7 (2). doi: 10.3390/app7020189

Lu, Y., Saeys, W., Kim, M., Peng, Y., and Lu, R. (2020). Hyperspectral imaging technology for quality and safety evaluation of horticultural products: A review and celebration of the past 20-year progress. *Postharvest Biol. Technol.* 170, 111318. doi: 10.1016/j.postharvbio.2020.111318

Magwaza, L. S., and Opara, U. L. (2015). Analytical methods for determination of sugars and sweetness of horticultural products-A review. *Scientia Hortic.* 184, 179–192. doi: 10.1016/j.scienta.2015.01.001

Magwaza, L. S., and Tesfay, S. Z. (2015). A review of destructive and non-destructive methods for determining avocado fruit maturity. *Food Bioprocess Technol.* 8 (10), 1995–20115. doi: 10.1007/s11947-015-1568-y

Mendy, T. K., Misran, A., Mahmud, T. M. M., and Ismail, S. I. (2019). Application of aloe vera coating delays ripening and extend the shelf life of papaya fruit. *Scientia Hortic.* 246, 769–776. doi: 10.1016/j.scienta.2018.11.054

Mesa, A. R., and Chiang, J. Y. (2021). Multi-input deep learning model with rgb and hyperspectral imaging for banana grading. *Agric. (Switzerland)* 11 (8), 1–18. doi: 10.3390/agriculture11080687

Metlenkin, D. A., Platov, Y. T., Platova, R. A., Zhirkova, E. V., and Teneva, O. T. (2022). Non-destructive identification of defects and classification of hass avocado fruits with the use of a hyperspectral image. *Agron. Res.* 20 (2), 326–340. doi: 10.15159/AR.22.027

Mishra, P., Woltering, E., and Harchioui, N. El (2020). Improved Prediction of 'Kent' Mango Firmness during Ripening by near-Infrared Spectroscopy Supported by Interval Partial Least Square Regression. *Infrared Phys. Technol.* 110, 103459. doi: 10.1016/j.infrared.2020.103459

Moreda, G. P., Ortiz-Cañavate, J., García-Ramos, F. J., and Ruiz-Altisent, M. (2009). Non-destructive technologies for fruit and vegetable size determination - A review. *J. Food Eng.* 92 (2), 119–136. doi: 10.1016/j.jfoodeng.2008.11.004

Morey, R., Ermolenkov, A., Payne, W. Z., Scheuring, D. C., Koym, J. W., Isabel Vales, M., et al. (2020). Non-invasive identification of potato varieties and prediction of the origin of tuber cultivation using spatially offset raman spectroscopy. *Analytical Bioanalytical Chem.* 412 (19), 4585–4945. doi: 10.1007/s00216-020-02706-5

Mukhametzyanov, R. R., Zaretskaya, A. S., Dzhancharova, G. K., Platonovskiy, N. G., and Ivantsova, N. N. (2022). "Russia as a subject of the world market for staple tropical fruits," in *Proceedings of the international scientific and practical conference strategy of development of regional ecosystems education-science-industry (ISPCR 2021, Springer Nature, Veliky Novgorod, Russia) 208 (Ispcr 2021)*, 594–602. doi: 10.2991/aebmr.k.220208.084

Ndlovu, P. F., Magwaza, L. S., Tesfay, S. Z., and Mphahlele, R. R. (2022). Destructive and rapid non-invasive methods used to detect adulteration of dried powdered horticultural products: A review. *Food Res. Int.* 157, 111198. doi: 10.1016/j.foodres.2022.111198

Neupane, C., Koirala, A., and Walsh, K. B. (2022). In-orchard sizing of mango fruit: 1. Comparison of machine vision based methods for on-the-go estimation. *Horticulturae* 8 (12), 1–17. doi: 10.3390/horticulturae8121223

Nicolaï, B. M., Beullens, K., Bobelyn, E., Peirs, A., Saeys, W., Theron, K. I., et al. (2007). Nondestructive measurement of fruit and vegetable quality by means of NIR spectroscopy: A review. *Postharvest Biol. Technol.* 46 (2), 99–1185. doi: 10.1016/j.postharvbio.2007.06.024

Nyarko, E. K., Vidović, I., Radočaj, K., and Cupec., R. (2018). A nearest neighbor approach for fruit recognition in RGB-D images based on detection of convex surfaces. *Expert Syst. Appl.* 114, 454–466. doi: 10.1016/j.eswa.2018.07.048

Okere, E. E., Arendse, E., Tsige, A. A., Perold, W. J., and Opara, U. L. (2022). Pomegranate quality evaluation using non-destructive approaches: A review. *Agric. (Switzerland)* 12 (12), 1–255. doi: 10.3390/agriculture12122034

Olarewaju, O. O., Bertling, I., and Magwaza, L. S. (2016). Non-Destructive evaluation of avocado fruit maturity using near infrared spectroscopy and PLS regression models. *Scientia Hortic.* 199, 229–236. doi: 10.1016/j.scienta.2015.12.047

Ozaki, Y. (2021). Infrared spectroscopy—Mid-infrared, near-infrared, and far-infrared/terahertz spectroscopy. *Analytical Sci.* 37 (9), 1193–1212. doi: 10.2116/analsci.20R008

Ozaki, Y., Christy, A. A., and McClure, W.F. (2006). *Near-infrared spectroscopy in food science and technology* (Hoboken, USA: John Wiley & Sons).

Pan, L., Sun, Ye, Xiao, H., Gu, X., Hu, P., Wei, Y., et al. (2017). Hyperspectral imaging with different illumination patterns for the hollowness classification of white radish. *Postharvest Biol. Technol.* 126, 40–49. doi: 10.1016/j.postharvbio.2016.12.006

Pandiselvam, R., Kaavya, R., Monteagudo, S. I., Divya, V., Jain, S., Khanashyam, A. C., et al. (2022). Contemporary developments and emerging trends in the application of spectroscopy techniques: A particular reference to coconut (Cocos nucifera L.). *Molecules* 27 (10), 1–22. doi: 10.3390/molecules27103250

Pathare, P. B., and Rahman, M. S. (2022). Nondestructive quality assessment techniques for fresh fruits and vegetables. In *Springer Nature*. doi: 10.1007/978-981-19-5422-1

Patrizi, B., De Cumis, M. S., Viciani, S., and D'Amato, F. (2019). Dioxin and related compound detection: perspectives for optical monitoring. *Int. J. Mol. Sci.* 20 (11), 2671. doi: 10.3390/ijms20112671

Pham, U. T., Phan, Q. H. T., Nguyen, L. P., Luu, P. D., Doan, T. D., Trinh, Ha T., et al. (2022). Rapid quantitative determination of multiple pesticide residues in mango fruits by surface-enhanced raman spectroscopy. *Processes* 10 (3), 1–14. doi: 10.3390/pr10030442

Phey, O., Hashim, N., and Maringgal, B. (2020). Quality evaluation of mango using non-destructive approaches: A review. *J. Agric. Food Eng.* 1 (1), 1–85. doi: 10.37865/jafe.2020.0003

Porat, R., Lichter, A., Terry, L. A., Harker, R., and Buzby, J. (2018). Postharvest losses of fruit and vegetables during retail and in consumers' Homes: quantifications, causes, and means of prevention. *Postharvest Biol. Technol.* 139, 135–149. doi: 10.1016/j.postharvbio.2017.11.019

Pratiwi, E. Z. D., Pahlawan, M. F.R., Rahmi, D. N., Amanah, H. Z., and Masithoh, R. E. (2023). Non-destructive evaluation of soluble solid content in fruits with various skin thicknesses using visible–shortwave near-infrared spectroscopy. *Open Agric.* 8 (1), 1–125. doi: 10.1515/opag-2022-0183

Pu, Y. Y., Feng, Y. Ze, and Sun, Da W. (2015). Recent progress of hyperspectral imaging on quality and safety inspection of fruits and vegetables: A review. *Compr. Rev. Food Sci. Food Saf.* 14 (2), 176–885. doi: 10.1111/1541-4337.12123

Purwanto, Y. A., Budiastra, I.W., Darmawati, E., and Arifiya, N. (2015). Measurement of starch and soluble solid content in papaya using near infrared spectroscopy. *Available Online Www.Jocpr.Com J. Chem. Pharm. Res.* 7 (6), 112–165.

Qin, J. (2010). "Hyperspectral imaging instruments," in *Hyperspectral imaging for food quality analysis and control* (England: Elsevier), 129–172.

Qin, (2012). "Hyperspectral and multispectral imaging in the food and beverage industries," in *Computer vision technology in the food and beverage industries* (Delhi: Elsevier), 27–63.

Qin, M. S.K., Chaoa, K., Dhakala, S., Chob, B.-K., Lohumib, C. M. S., Pengd, Y., et al. (2019). Advances in raman spectroscopy and imaging techniques for quality and safety inspection of horticultural products. *Postharvest Biol. Technol.* 149, 101–117. doi: 10.1016/j.postharvbio.2018.11.004

Qin, K. C., Kim, M. S., Lu, R., and Burks, T. F. (2013). Hyperspectral and multispectral imaging for evaluating food safety and quality. *J. Food Eng.* 118 (2), 157–715. doi: 10.1016/j.jfoodeng.2013.04.001

Rahman, A., and Cho, B. K. (2016). Assessment of seed quality using non-Destructive measurement techniques: A review. *Seed Sci. Res.* 26 (4), 285–3055. doi: 10.1017/S0960258516000234

Raj, T. S., and Suji, H. A. (2019). Post-harvest quality of fresh produce. *Adv. Agri. Sci.* 129–143. doi: 10.22271/ed.book21

Rajkumar, P., Wang, N., EImasry, G., Raghavan, G. S. V., and Gariepy, Y. (2012). Studies on banana fruit quality and maturity stages using hyperspectral imaging. *J. Food Eng.* 108 (1), 194–200. doi: 10.1016/j.jfoodeng.2011.05.002

Retamales, J. B. (2011). World Temperate Fruit Production: Characteristics and Challenges | Produção Mundial de Frutas de Clima Temperado : Caracteristicas e Desafios. *Rev. Bras. Fruticultura 33 (SPEC. ISSU)* 33, 121–130. doi: 10.1590/S0100-29452011000500015

Rivera, N. Ve´lez, Gómez-Sanchis, J., Chanona-Pérez, J., Carrasco, J. José, Millán-Giraldo, Mónica, Lorente, D., et al. (2014). Early detection of mechanical damage in mango using NIR hyperspectral images and machine learning. *Biosyst. Eng.* 122, 91–98. doi: 10.1016/j.biosystemseng.2014.03.009

Rostron, P., Gaber, S., and Gaber, D. (2016). Raman Spectroscopy, Review. *Int. J. Eng. Tech. Res. (IJETR)* 6, 2454–4698.

Ruiz-Altisent, M., Ruiz-Garcia, L., Moreda, G. P., Lu, R., Hernandez-Sanchez, N., Correa, E. C., et al. (2010). Sensors for product characterization and quality of specialty crops-A review. *Comput. Electron. Agric.* 74 (2), 176–945. doi: 10.1016/j.compag.2010.07.002

Rungpichayapichet, P., Mahayothee, B., Nagle, M., Khuwijitjaru, P., and Müller, J. (2016). Robust NIRS models for non-destructive prediction of postharvest fruit ripeness and quality in mango. *Postharvest Biol. Technol.* 111, 31–40. doi: 10.1016/j.postharvbio.2015.07.006

Ryu, M., Ng, S. H., Anand, V., Lundgaard, S., Hu, J., Katkus, T., et al. (2021). Attenuated total reflection at THz wavelengths: prospective use of total internal reflection and polariscopy. *Appl. Sci.* 11 (16), 76325. doi: 10.3390/app11167632

Sahu, D., and Potdar, R. M. (2017). Defect identification and maturity detection of mango fruits using image analysis. *Am. J. Artif. Intell.* 1 (1), 5–145. doi: 10.11648/j.ajai.20170101.12

Salehi, F. (2020). Recent advances in the modeling and predicting quality parameters of fruits and vegetables during postharvest storage: A review. *Int. J. Fruit Sci.* 20 (3), 506–520. doi: 10.1080/15538362.2019.1653810

Sanchez, P. D. C., Hashim, N., Shamsudin, R., and Nor, M. Z. M. (2020). Applications of imaging and spectroscopy techniques for non-destructive quality evaluation of potatoes and sweet potatoes: A review. *Trends Food Sci. Technol.* 96, 208–221. doi: 10.1016/j.tifs.2019.12.027

Sarkar, T., Chandra, B., Viswavidyalaya, K., and Mani, A. (2018) *Maturity indices of tropical and sub-tropical fruit crops 38 MATURITY INDICES OF TROPICAL AND SUB-TROPICAL FRUIT CROPS*. Available at: https://www.researchgate.net/publication/329266894.

Sebben, J. Antônio, Espindola, J. da S., Ranzan, L., Moura, N. F. de, Trierweiler, L. F., and Trierweiler, J. Otávio (2018). Development of a quantitative approach using raman spectroscopy for carotenoids determination in processed sweet potato. *Food Chem.* 245, 12–31. doi: 10.1016/j.foodchem.2017.11.086

Shewfelt, R. L. (2014). "Measuring quality and maturity," in *Postharvest handling* (New York: Elsevier), 387–410.

Si, W., Xiong, J., Huang, Y., Jiang, X., and Hu, D. (2022). Quality assessment of fruits and vegetables based on spatially resolved spectroscopy: A review. *Foods* 11 (9), 1–215. doi: 10.3390/foods11091198

Siesler, H. W., Kawata, S., Michael Heise, H., and Ozaki, Y. (2008). *Near-infrared spectroscopy: principles, instruments, applications* (Weinheim, German: John Wiley & Sons).

Silva, C. E. de F., and Abud, A. K. de S. (2017). Tropical fruit pulps: processing, product standardization and main control parameters for quality assurance. *Braz. Arch. Biol. Technol.* 60, 1–19. doi: 10.1590/1678-4324-2017160209

Sinanoglou, V. J., Tsiaka, T., Aouant, K., Mouka, E., Ladika, G., Kritsi, E., et al. (2023). Quality assessment of banana ripening stages by combining analytical methods and image analysis. *Applied Sciences (Switzerland)* 13 (6), 3533.

Sirisomboon, P. (2018). NIR spectroscopy for quality evaluation of fruits and vegetables. *Materials Today: Proc.* 5 (10), 22481–22486. doi: 10.1016/j.matpr.2018.06.619

Sohaib, A. Z., Qureshi, W. S., Arslan, M., Malik, A. U., Alasmary, W., and Alanazi., E. (2020). Towards fruit maturity estimation using NIR spectroscopy. *Infrared Phys. Technology.* 111, 1–17. doi: 10.1016/j.infrared.2020.103479

Srivichien, S., Terdwongworakul, A., and Teerachaichayut, S. (2015). Quantitative prediction of nitrate level in intact pineapple using vis-NIRS. *J. Food Eng.* 150, 29–34. doi: 10.1016/j.jfoodeng.2014.11.004

Su, W. H., Bakalis, S., and Sun, Da W. (2019). Chemometrics in tandem with near infrared (NIR) hyperspectral imaging and fourier transform mid infrared (FT-MIR) microspectroscopy for variety identification and cooking loss determination of sweet potato. *Biosyst. Eng.* 180, 70–86. doi: 10.1016/j.biosystemseng.2019.01.005

Su, W. H., and Sun, Da W. (2018). Fourier transform infrared and raman and hyperspectral imaging techniques for quality determinations of powdery foods: A review. *Compr. Rev. Food Sci. Food Saf.* 17 (1), 104–225. doi: 10.1111/1541-4337.12314

Su, W.-H., and Sun, D.-W. (2019). Rapid determination of starch content of potato and sweet potato by using NIR hyperspectral imaging. *Hortscience* 54, S38.

Subedi, P. P., and Walsh, K. B. (2011). Assessment of sugar and starch in intact banana and mango fruit by SWNIR spectroscopy. *Postharvest Biol. Technol.* 62 (3), 238–245. doi: 10.1016/j.postharvbio.2011.06.014

Subedi, P. P., and Walsh, K. B. (2020). Assessment of avocado fruit dry matter content using portable near infrared spectroscopy: method and instrumentation optimisation. *Postharvest Biol. Technol.* 161, 1–10. doi: 10.1016/j.postharvbio.2019.111078

Tang, T., Zhang, M., and Mujumdar, A. S. (2022). Intelligent detection for fresh-cut fruit and vegetable processing: imaging technology. *Compr. Rev. Food Sci. Food Saf.* 21 (6), 5171–5985. doi: 10.1111/1541-4337.13039

Tata, J. S., Kalidindi, N. K. V., Katherapaka, H., Julakal, S. K., and Banothu., M. (2022). Real-time quality assurance of fruits and vegetables with artificial intelligence. *J. Physics: Conf. Ser.* 2325 (1), 1–13. doi: 10.1088/1742-6596/2325/1/012055

Torres-Sánchez, R., Martínez-Zafra, MaríaT., Castillejo, N., Guillamón-Frutos, A., and Artés-Hernández, F. (2020). Real-time monitoring system for shelf life estimation of fruit and vegetables. *Sensors (Switzerland)* 20 (7), 1–21. doi: 10.3390/s20071860

Tsuchikawa, S., Ma, Te, and Inagaki, T. (2022). Application of near-infrared spectroscopy to agriculture and forestry. *Analytical Sci.* 38 (4), 635–425. doi: 10.1007/s44211-022-00106-6

Uarrota, V. G., and Pedreschi, R. (2022). Mathematical Modelling of Hass Avocado Firmness by Using Destructive and Non-Destructive Devices at Different Maturity Stages and under Two Storage Conditions. *Folia Horticulturae.* 34 (2), 139–150. doi: 10.2478/fhort-2022-0011

Ucat, R. C., and Dela Cruz, J. C. (2019). "Postharvest grading classification of cavendish banana using deep learning and tensorflow," in *2019 international symposium on multimedia and communication technology, ISMAC 2019* 1, 6. doi: 10.1109/ISMAC.2019.8836129

Wang, K., Li, Z., Li, J., and Lin, H. (2021). Raman spectroscopic techniques for nondestructive analysis of agri-foods: A state-of-the-art review. *Trends Food Sci. Technol.* 118, 490–504. doi: 10.1016/j.tifs.2021.10.010

Wang, Z., Walsh, K. B., and Verma, B. (2017). On-tree mango fruit size estimation using RGB-D images. *Sensors (Switzerland)* 17 (12), 1–155. doi: 10.3390/s17122738

Wang, A., Hu, D., and Xie, L. (2014). Comparison of detection modes in terms of the necessity of visible region (VIS) and influence of the peel on soluble solids content (SSC) determination of navel orange using VIS-SWNIR spectroscopy. *J. Food Eng.* 126, 126–132. doi: 10.1016/j.jfoodeng.2013.11.011

Wang, M. Hu, and Zhai, G. (2018). Application of deep learning architectures for accurate and rapid detection of internal mechanical damage of blueberry using hyperspectral transmittance data. *Sensors (Switzerland)* 18 (4), 1–145. doi: 10.3390/s18041126

Wu, Di, and Sun, Da W. (2013). Advanced applications of hyperspectral imaging technology for food quality and safety analysis and assessment: A review - part I: fundamentals. *Innovative Food Sci. Emerging Technol.* 19, 1–14. doi: 10.1016/j.ifset.2013.04.014

Xu, H., Ren, J., Lin, J., Mao, S., Xu, Z., Chen, Z., et al. (2023). The impact of high-Quality data on the assessment results of visible/near-Infrared hyperspectral imaging and development direction in the food fields: A review *J. Food Measurement Characterization* 17 (3), 2988–3004. doi: 10.1007/s11694-023-01822-x

Yahaya,, Mardziah, O. K., and Omar, A. F. (2017). *Spectroscopy of tropical fruits: sala mango and B10 carambola (Penerbit USM)* (Penang, Malaysia: Penerbit USM).

Yahaya, O. K. M., Matjafri, M. Z., Aziz, A. A., and Omar, A. F. (2011). Non-destructive quality evaluation of fruit by color based on RGB LEDs system. *2014 2nd Int. Conf. Electronic Design ICED 2014* 1001, 230–233. doi: 10.1109/ICED.2014.7015804

Yang,, and Xu, Y. (2021). Applications of deep-learning approaches in horticultural research: A review. *Horticulture Res.* 8 (1), 1–31. doi: 10.1038/s41438-021-00560-9

Yang, J., Yin, C., Miao, Xu, Meng, X., Liu, Z., and Hu, L. (2021). Rapid discrimination of adulteration in radix astragali combining diffuse reflectance mid-infrared fourier transform spectroscopy with chemometrics. *Spectrochimica Acta - Part A: Mol. Biomolecular Spectrosc.* 248, 119251. doi: 10.1016/j.saa.2020.119251

Ye, D., Sun, L., Tan, W., Che, W., and Yang, M. (2018). Detecting and classifying minor bruised potato based on hyperspectral imaging. *Chemometrics Intelligent Lab. Syst.* 177, 129–139. doi: 10.1016/j.chemolab.2018.04.002

Yeap, K. Ho, and Hirasawa, K. (2019). Introductory chapter: electromagnetism. *Electromagnetic Fields Waves* 356, 3–10. doi: 10.5772/intechopen.85155

Zainalabidin, F. A., Sagrin, M. S., Azmi, W. N. W., and Ghazali, A. S. (2019). Optimum postharvest handling-effect of temperature on quality and shelf life of tropical fruits and vegetables. *J. Trop. Resour. Sustain. Sci. (JTRSS)* 7 (1), 23–305. doi: 10.47253/jtrss.v7i1.505

Zakaria, L. (2023). Fusarium species associated with diseases of major tropical fruit crops. *Horticulturae* 9 (3), 322. doi: 10.3390/horticulturae9030322

Zhang, B., Huang, W., Li, J., Zhao, C., Fan, S., Wu, J., et al. (2014). Principles, developments and applications of computer vision for external quality inspection of fruits and vegetables: A review. *Food Res. Int.* 62, 326–343. doi: 10.1016/j.foodres.2014.03.012

Zhang, L., Huang, Y., Sun, F., Chen, Da, Netzel, M., Heather, E., et al. (2021). The effect of maturity and tissue on the ability of mid infrared spectroscopy to predict the geographical origin of banana (Musa cavendish). *Int. J. Food Sci. Technol.* 56 (6), 2621–2275. doi: 10.1111/ijfs.14960

Zhang, J. Y., Lin, T., and Ying, Y. (2021). Food and agro-product quality evaluation based on spectroscopy and deep learning: A review. *Trends Food Sci. Technol.* 112, 431–441. doi: 10.1016/j.tifs.2021.04.008

Zhu, D., Ren, X., Wei, L., Cao, X., Ge, Y., Liu, He, et al. (2020). Collaborative analysis on difference of apple fruits flavour using electronic nose and electronic tongue. *Scientia Hortic.* 260, 108879. doi: 10.1016/j.scienta.2019.108879

frontiers | Frontiers in Plant Science

TYPE Original Research
PUBLISHED 24 August 2023
DOI 10.3389/fpls.2023.1249989

OPEN ACCESS

EDITED BY
Xi Tian,
Beijing Academy of Agriculture and Forestry Sciences, China

REVIEWED BY
Yaoyao Fan,
Shenyang Agricultural University, China
Guoxiong Zhou,
Central South University Forestry and Technology, China
Yu Xia,
Shaanxi University of Science and Technology, China

*CORRESPONDENCE
Xiaofei Fan
leopardfxf@163.com

†These authors have contributed equally to this work

RECEIVED 29 June 2023
ACCEPTED 04 August 2023
PUBLISHED 24 August 2023

CITATION
Song C, Peng B, Wang H, Zhou Y, Sun L, Suo X and Fan X (2023) Maize seed appearance quality assessment based on improved Inception-ResNet. *Front. Plant Sci.* 14:1249989. doi: 10.3389/fpls.2023.1249989

Maize seed appearance quality assessment based on improved Inception-ResNet

Chang Song†, Bo Peng†, Huanyue Wang, Yuhong Zhou, Lei Sun, Xuesong Suo and Xiaofei Fan*

College of Mechanical and Electrical Engineering, Hebei Agricultural University, Baoding, China

Current inspections of seed appearance quality are mainly performed manually, which is time-consuming, tedious, and subjective, and creates difficulties in meeting the needs of practical applications. For rapid and accurate identification of seeds based on appearance quality, this study proposed a seed-quality evaluation method that used an improved Inception-ResNet network with corn seeds of different qualities. First, images of multiple corn seeds were segmented to build a single seed image database. Second, the standard convolution of the Inception-ResNet module was replaced by a depthwise separable convolution to reduce the number of model parameters and computational complexity of the network. In addition, an attention mechanism was applied to improve the feature learning performance of the network model and extract the best image information to express the appearance quality. Finally, the feature fusion strategy was used to fuse the feature information at different levels to prevent the loss of important information. The results showed that the proposed method had decent comprehensive performance in detection of corn seed appearance quality, with an average of 96.03% for detection accuracy, 96.27% for precision, 96.03% for recall rate, 96.15% for F1 value of reconciliation, and the average detection time of an image was about 2.44 seconds. This study realized rapid nondestructive detection of seeds and provided a theoretical basis and technical support for construction of intelligent seed sorting equipment.

KEYWORDS

corn seed, quality assessment, depthwise separable convolution, attention mechanism, feature fusion

1 Introduction

Maize is an essential cereal crop that is widely grown worldwide and has an increasing production and trade volume (Ali et al., 2020). Appearance quality is an important factor affecting the price of corn seeds, and effective identification of seed quality is critical for ensuring food security and agricultural production safety. With the rapid advancements in automation, machine vision technology (Huang et al., 2019; Kim et al., 2020; Wang and Xiao, 2020; Ansari et al., 2021; Lu et al., 2022) can be used to nondestructively and quickly

obtain seed surface feature information at a low cost and high detection accuracy and efficiency, thereby providing potential new methods for seed quality identification.

Machine learning (Rahmani et al., 2016; Prajapati et al., 2018; Liang et al., 2019; Sharaff et al., 2021) is an active field of artificial intelligence research that has advantages in terms of training small data samples and wide applications in agricultural product identification and defect detection. Farooqui et al. (2019) used a gray-level co-occurrence matrix for disease feature extraction, a support vector machine classifier for plant disease identification, combined with advanced neural network to optimize the data to improve the detection accuracy and demonstrated the feasibility of this approach for plant disease diagnosis through experiments. Effective classification of seeds is an important part of selecting and breeding good seeds. To simplify the seed selection process, Ma et al. (2021) proposed a peanut seed appearance quality detection method, using peanut seed size and appearance color as the main features and a support vector machine classification model to complete the classification task. The experimental results showed that the method had an accuracy of 86% for the classification of bulk peanut seeds, which met the preliminary classification requirements of peanut seeds in actual production. Gao et al. (2016) designed a fresh corn quality detection classifier, which analyzed the texture features of fresh corn images by wavelet analysis method, used the maximum entropy function to measure the separation degree of the texture images, and combined with the weight criterion to classify the fresh corn of different varieties, sizes, and damage degrees, and the effective classification rate could reach more than 99%. Zhao et al. (2022) extracted three categories of raw coffee bean features: contour, color, and texture to detect defective raw coffee beans by the features of a single category or category combinations. The findings were applied to a grid search to determine support vector machine classification model parameters and combined with a k-fold cross-validation test to compare support vector machine model performances. The experimental results showed that the average accuracy, precision, recall, and F1 values were 84.9%, 85.8%, 82.3%, and 84.0%, respectively. This method provided a theoretical base for the automatic detection of defective raw coffee beans. With the requirements of strict and precise agricultural development, there is an urgent need to explore new research methods to achieve precise assessment of seed appearance quality and promote intelligent agricultural development.

With the rapid development of deep learning, convolutional neural networks are widely used in the fields of medicine, aviation, and agriculture because of their excellent feature learning and expression capabilities (Kamilaris and Prenafeta-Boldu, 2018; Naranjo-Torres et al.,2020; Zhang et al., 2020; Cong and Zhou, 2022; Liu et al., 2023). Compared with traditional machine learning techniques, convolutional neural networks are more generalizable, faster to train, and can obtain significant information directly from images, which eliminates the tedious steps of manually extracting image features used in traditional methods. In applications for agriculture, convolutional neural networks are often used in areas such as the classification of crop pests and diseases (Wu et al., 2019; Peng et al., 2019; Tiwari et al., 2021; Liu et al., 2022; Liu et al., 2022), agricultural product species identification (Ajit et al., 2020; Gao et al., 2020; Chen et al., 2021; Laabassi et al., 2021; Sj et al.,2021), yield estimation (Zhang et al., 2020; Tan et al., 2019; Alexandros et al., 2023; Kavita et al., 2023), and crop quality grading (Anikó and Miklós, 2022; Liu et al., 2022; Li et al., 2022; Wang Z. et al., 2022; Peng et al., 2023), in which they greatly promote the development of agricultural intelligence. Along with the arrival of the era of big data, the amount of image information increases exponentially, resulting in an increase in the amount of computation and training difficulty in the training process. This has also led researchers to pay more attention to lightweight networks, in order to maintain the accuracy of the premise of lightweight transformation of the network, MobileNet, ShuffleNet and other lightweight networks came into being, which can be better adapted to the evolving needs of the mobile market. To effectively alleviate the large amounts of computational resources and storage costs required for real-time image processing, Yuan et al. (2022) constructed a high-performance low-resolution MobileNet model, in which the network structure was simplified by cropping and the inception structure was used to fill the Dwise layer in a depth-separable convolution to extract the richer low-resolution features. The experimental results showed that the model achieved 89.38%, 71.60%, and 87.08% accuracies with the CIFAR - 10, CIFAR - 100, and CINIC - 10 datasets, respectively, and was suitable for real-time image classification tasks in low-resolution application scenarios. Fang et al. (2022) proposed a new network structure, HCA-MFFNet, for maize leaf disease recognition in complex contexts, and in order to validate the feasibility and effectiveness of the model in complex environments, it was compared with the existing methods, and the results proved that the model had an average recognition accuracy of 97.75% and an F1 value of 97.03%, which was the best overall performance. Hou et al. (2020) proposed a damage classification algorithm for castor seeds based on a convolutional neural network. Authors used castor seeds with missing shells or cracks and intact castor seeds to construct a dataset and build a network model to classify the seeds. The experimental results showed that the average accuracy was 87.78%, with 96.67% for castor seeds without shells, 80.00% for cracked castor seeds, and 86.67% for intact castor seeds; therefore, this method provided a feasible solution for the online real-time classification of castor seeds. Wang L. et al. (2022) designed a defect detection method based on the watershed algorithm and a two-channel convolutional neural network model, which can effectively identify defective and non-defective seeds with an average accuracy of 95.63%, an average recall of 95.29% and an F1 value of 95.46%. The assay provided an effective tool for the detection of corn seed defects. Cai et al. (2023) proposed a new grape leaf disease identification model, which was proved to have an identification accuracy of 93.26%, effectively providing decision-making information for the grape leaf disease identification system in precision agriculture. The above research results simplify the network structure complexity to a certain extent and reduced the requirement of hardware devices for model training. However, this is also often accompanied by poor model recognition accuracy, making it difficult to meet the needs of practical applications. Therefore, additional methods must be developed to improve the accuracy and achieve the purpose of accurate recognition.

In this study a seed-appearance quality assessment method based on an improved Inception-ResNet was proposed using intact and damaged maize seeds as test samples. By replacing depthwise separable convolution, adding an attention mechanism, and introducing a feature fusion strategy, the Inception-ResNet network structure was improved and optimized to obtain more detailed feature information, with the aims of achieving accurate, rapid, and nondestructive detections of seed appearance quality and providing a feasible reference scheme for subsequent automatic seed quality sorting processes.

2 Materials and methods

2.1 Experimental materials and treatment

2.1.1 Image acquisition

In this experiment, 50 groups of corn seeds with good appearance and broken appearance were collected respectively, with a total of 982 seeds, including 458 good seeds and 524 broken seeds. The training set and test set were divided according to 4:1. Ten sets of maize seed image data containing both good and defective seeds were used for verification of the final model (Figure 1). The image acquisition platform mainly consisted of four parts: a multispectral surface array camera from JAI Company in Denmark, with an image resolution of 2048×1536; a bracket to adjust the camera height; a light source on both sides of the camera; and a shelf to place the maize seed samples. The image data acquisition device is shown in Figure 2.

2.1.2 Image processing

The Python3.6 script language was used to segment the corn seed images, as shown in Figure 3A. First, the original color image was converted to grayscale, and a binary image was then obtained using the adaptive thresholding method, whereby the seed region is shown as white and the background region as black. White noise in the image was removed using a morphological open operation, and expansion was used in the foreground to distinguish the background and foreground areas of the image. The distanceTransform function was then used to obtain the center region of each seed, and the expanded image was subtracted from the central region to obtain the edge region. Finally, the watershed algorithm was used to mark the identified central region, delineate the seed boundary, and determine the range of each seed in the complete image by the location coordinates for segmentation. The segmentation effect is shown in Figure 3B.

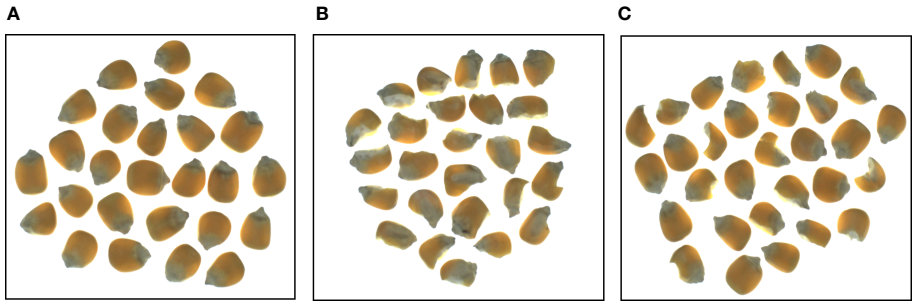


FIGURE 1
Corn seeds with different appearance qualities **(A)** Good seed grain, **(B)** Defective grain, **(C)** Both conditions.

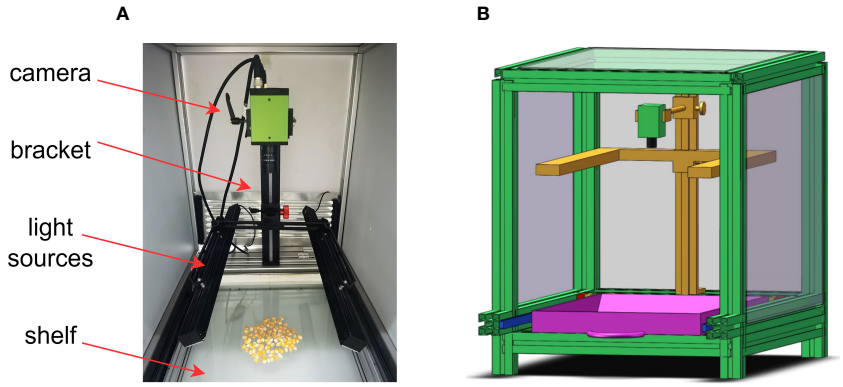


FIGURE 2
Image acquisition platform **(A)** physical device diagram and **(B)** 3D view of the device.

10.3389/fpls.2023.1249989

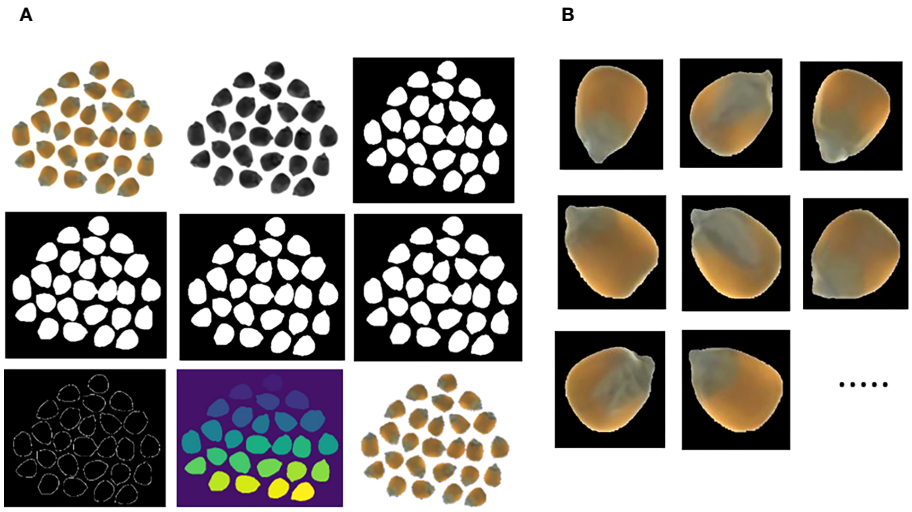


FIGURE 3
Process of watershed algorithm **(A)** Segmentation process and **(B)** Segmentation results.

2.2 Basic method and test environment

2.2.1 Convolutional neural networks

A convolutional neural network (Rong et al., 2019) is a kind of multilayer perceptron. A traditional convolutional neural network consists of an input layer, convolutional layer, pooling layer, and fully connected layer. A simple neural network model can be formed by mixing different depths and stacking orders, as shown in Figure 4, in which the term random represents the number of times a particular layer of a structure is randomized.

The input layer represents the initial input of the entire convolutional neural network, and in the case of image processing, the input layer contains the pixel matrix of the input image, and the convolutional layer is the core component of convolutional neural network, which has the characteristics of local connection, weight sharing and translation invariance, and realizes the feature extraction function of the network. The pooling layer compresses the size of the input feature map while extracting the main features, thereby reducing the complexity of the network training process and the probability of overfitting the network model during this process. The fully connected layer is involved in weighting the features extracted from the convolutional and pooling layers, transforming them into layer vectors, and inputting one-dimensional data into the Softmax layer through multiplication operations to obtain the image classification results. It can be understood as a parallel, large-scale distributed processor that is capable of storing and using empirical knowledge.

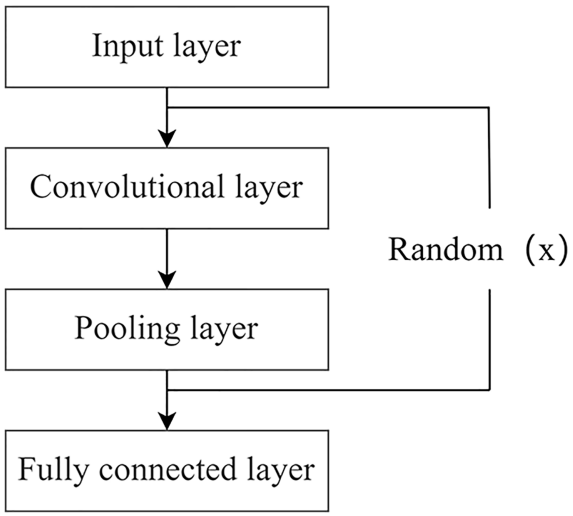


FIGURE 4
Basic structure of convolutional neural network.

2.2.2 Depthwise separable convolution

Depthwise separable convolution (Mamalet and Garcia, 2012) was proposed as a typical lightweight convolution structure that has a significantly reduced parameter number and increased training speed compared to a standard convolution and can separate channels and regions in the convolution operation. As shown in Figure 5, depthwise separable convolution mainly consists of two parts, depthwise convolution and pointwise convolution, which are used to extract feature information. One depthwise convolution kernel is responsible for one channel, and the number of feature maps generated by this process equals the number of input channels. The convolution operation is performed independently for each channel of the input layer, which does not make full use of the feature information from other channels in the same spatial location. Therefore, pointwise convolution is required to combine the feature maps to generate a new feature map, in which the number of convolution kernels corresponds to the output feature map.

Depthwise separable convolution effectively reduces the number of parameters required for the network model compared to the normal convolutional approach. The N H × W × C convolution kernels can be replaced by C H × W × 1 depthwise and N 1 × 1 × C pointwise convolution kernels. The number of depthwise convolution parameters is (H × W × 1) × C, and the number of pointwise convolution parameters is (1 × 1 × C) × N. The combined number of parameters for the depthwise separable convolution can be calculated as follows:

$$Params = H \times \mathrm{W} \times \mathrm{C} + \mathrm{C} \times \mathrm{N} \tag{1}$$

The number of parameters in the ordinary convolution is H × W × C × N, and the relationship between the two is compared as follows:

$$\frac{H \times \mathrm{W} \times \mathrm{C} + \mathrm{C} \times \mathrm{N}}{H \times \mathrm{W} \times \mathrm{C} \times \mathrm{N}} = \frac{1}{N} + \frac{1}{H \times \mathrm{W}} \tag{2}$$

2.2.3 Attention mechanism

The working principle of the attention mechanism is similar to the selective attention of human vision. It utilizes limited computational resources to focus on important feature information and emphasize regions of interest in a dynamically weighted manner, discarding irrelevant background information and nuisance information in the input features to improve network performance. In general, the attention mechanism involves the process of weight assignment, in which the input feature information is processed, the attention information is obtained through a weight assignment, and the attention mechanism is used to process these features. The process can be expressed as

$$Attention = f(g(x), x) \tag{3}$$

Here, x represents the input feature information of the attention mechanism; g(x) represents the focus on the key areas, that is, the process of generating attention information by processing the input features; and f(g(x), x) represents the processing of the key areas, that is, processing of the input information using the attention information obtained from g(x).

2.2.4 Evaluation indicators

In this study, objective evaluation criteria (Feng et al., 2022; Wang X. et al., 2022) were used to analyze the inspection model of the maize seed appearance quality using Accuracy(A), Precision (P), Recall (R), and by introducing the F1 value as the average evaluation of the reconciliation. The related formulae are as follows:

$$A = \frac{\mathrm{TP} + \mathrm{TN}}{\mathrm{TP} + \mathrm{TN} + \mathrm{FN} + \mathrm{FP}} \times 100\% \tag{4}$$

$$P = \frac{\mathrm{TP}}{\mathrm{TP} + \mathrm{FP}} \times 100\,\% \tag{5}$$

$$R = \frac{\mathrm{TP}}{\mathrm{TP} + \mathrm{FN}} \times 100\,\% \tag{6}$$

$$F1 = \frac{2PR}{P + R} \times 100\,\% \tag{7}$$

2.2.5 Test environment

We used a Windows 10, 64-bit operating system with a x64 based processor, Cuda version 11.0, and the Tensorflow deep learning framework based on the Python programming language. The computer contained an NVIDIA GeForce RTX 3090 graphics card with 24G video memory and a 12th Gen Intel(R) Core (TM) i7-12700KF processor at 3.61 GHz.

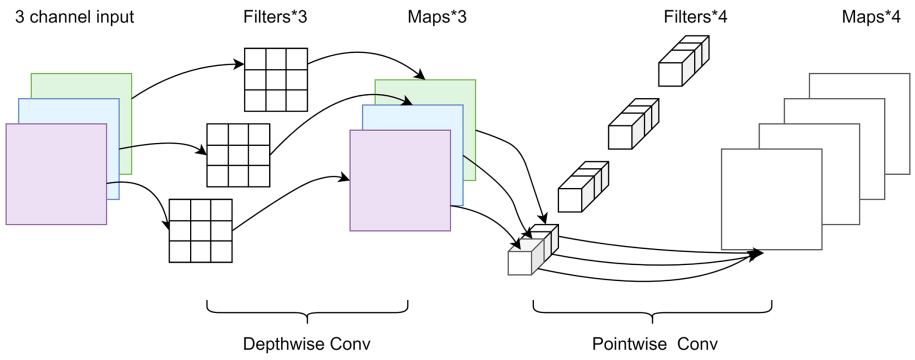


FIGURE 5
Depthwise separable convolution.

2.3 Algorithm improvement

2.3.1 Improving the Inception-ResNet module

The traditional Inception module, shown in Figure 6A, consists of 1 × 1, 3 × 3, and 5 × 5 convolution operations of various sizes for feature extraction using multiple-scale parallel convolution operations. The Inception-ResNet module (Figure 6B) combined the residual network structure of the ResNet with the separation of the large convolutional network into two tandem small convolutional structures to obtain the output feature maps of the 5 × 5 convolution, thereby improving the classification performance of the model. In this study, based on the Inception-ResNet module, the standard convolution part of the Inception-ResNet structure was replaced by a cost-effective convolution operation by combining the lightweight structure depthwise separable convolution (Dinception-ResNet). This approach could reduce the number of model parameters, increase the depth of the network, and enhance the feature extraction capability of the model, while preserving the feature diversity of the traditional inception multigroup structure.

A comparison of the parameters of the improved model is presented in Table 1. The original Inception-ResNet model used up to 28 979 618 parameters; this number was reduced by 177 920 after the depthwise separable convolutional replacement and the ratio of the number of trainable parameters to the total number of parameters was improved, which favors the design of lightweight networks.

2.3.2 Adding attention mechanism

Introducing the Efficient Channel Attention Network after the Dinception-ResNet module effectively avoided dimensionality reduction and captured cross-channel interactions, as shown in Figure 7. The Efficient Channel Attention Network first transformed the input feature map from a matrix [H, W, C] into a vector [1, 1, C] using global averaging pooling, after which it calculated the adaptive 1D convolution kernel size based on the number of channels of the feature map, which was then used in 1D convolution. Subsequently, the weights of the feature maps with respect to each channel were obtained. Finally, the normalized weights and the original input feature maps were multiplied channel-by-channel to generate the weighted feature maps. These maps can be used to solve the information overload problem and improve the efficiency and accuracy of task processing by focusing on the more critical information for the current task, reducing the attention to other information, and using a small number of parameters to achieve suitable results. To fully demonstrate the effectiveness of Efficient Channel Attention Network, we conducted a comparison experiment between it and the commonly used attention mechanisms CBAM, SENet and CANet. By testing the relevant indexes of the experiments, the results were shown in Table 2, which clearly concluded that the overall detection performance of Efficient Channel Attention Network was optimal.

2.3.3 Introducing the feature fusion mechanism

As the depth of the network model increases, important feature information may be missed in the feature extraction process,

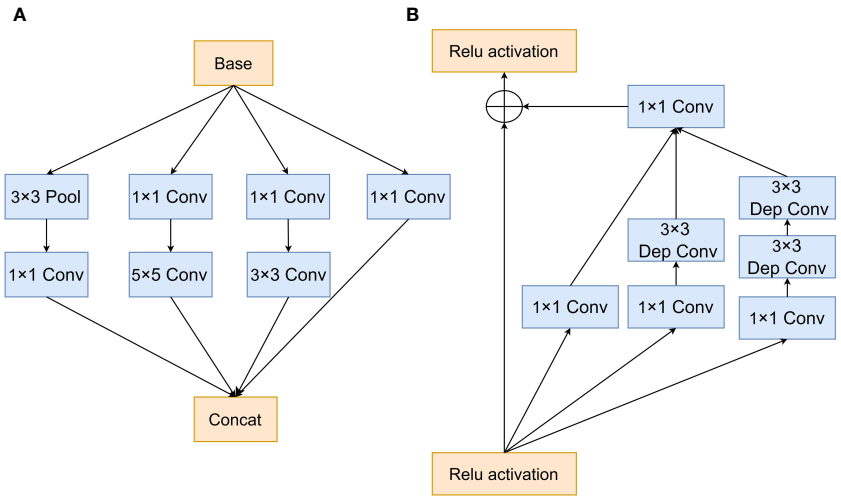


FIGURE 6
Inception module **(A)** Traditional Inception module and **(B)** Dinception-ResNet module.

TABLE 1 Parameter comparison.

Model	Total parameters	Trainable parameters	Non-trainable parameters
Inception-ResNet	28 979 618	28 947 106	32 512
Dinception-ResNet	28 801 698	28 770 338	31 360

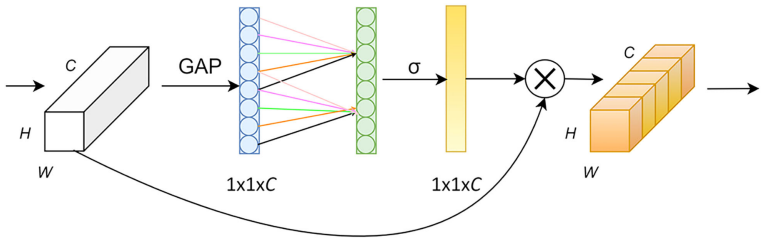


FIGURE 7
ECA network structure.

thereby decreasing the classification accuracy of the model. Therefore, this study proposed feature fusion, as shown in Figure 8, in which the output features of each module were first pooled by global averaging to avoid overfitting. Subsequently, the output feature maps of the three stages were feature-fused to enrich the output feature information, enhance the generalization performance of the model, and improve the expression ability of the feature information.

2.3.4 Improved network model structure

The traditional convolutional neural network structure generates a large number of parameters in the deep network. In addition, key features are mixed with irrelevant features, and the network is difficult to optimize; therefore, this study proposed an improved model based on Inception-ResNet (Figure 9). Depthwise separable convolution was used to replace the standard convolution in the Inception-ResNet module to reduce the number of model parameters, the network model was optimized by introducing the Efficient Channel Attention mechanism to increase the feature weights of key information and improve the network performance; in addition, the output features of the low, middle and high layers of the model were fused to improve the feature extraction ability of the model and enrich feature information to achieve the purpose of network optimization.

2.3.5 Ablation study

In order to validate the effectiveness of the method proposed in this paper, separate experiments were conducted for the proposed modules to compare with the original Inception-Resnet model. By comparing the results of groups 1-2-3-4 in Table 3, it can be seen that the performance of each module reference is improved to some extent compared to the original network. Among them, the introduction of the feature fusion module resulted in the most obvious model performance improvement, with each performance index improved by 3.38 – 5.55 percentage points respectively. The average time for detecting an image was 2.62 seconds. Replacing standard convolution with depthwise separable convolution, the average time to detect an image was reduced by 30 milliseconds. By comparing the results of 1-2-5-6 groups and referring modules one by one, the model performance was improved and the effectiveness of the modules is fully proved.

3 Results and analysis

3.1 Model optimization

Different networks apply different model parameters. By adjusting certain parameters in the model and using test set recognition accuracy as the evaluation index, we investigated the effects of the parameters on the classification accuracy of the model. In the convolutional neural network model, batch size is an important hyperparameter, and we chose a batch size that was appropriate to train the model to converge to the global optimum. Using a large number of parameters in a deep convolutional neural network creates correction issues; therefore, a suitable optimizer was selected to improve the model training speed and accuracy. The training parameter batch sizes of the model were 16, 32, and 64, and two different optimization algorithms, Adam and SGD, were used. The number of model iterations was set to 100 and the learning rate was 0.001; the final test set accuracy variation curve is shown in Figure 10. These settings ensured that when the Adam optimizer was selected for the improved model and the bath size was set to 32, the accuracy of the test set was the highest and the detection performance was optimal.

TABLE 2 Comparison of detection performance of different attention mechanisms.

Attention mechanism	Average Accuracy/%	Average Precision/%	Average Recall/%	F1/%	Detection time/s
ECANet	95.96	92.46	91.26	91.86	2.68
CBAM	89.68	91.33	89.68	90.50	2.65
SENet	83.33	87.34	83.33	85.29	2.67
CANet	90.48	91.89	90.47	91.17	2.66

10.3389/fpls.2023.1249989

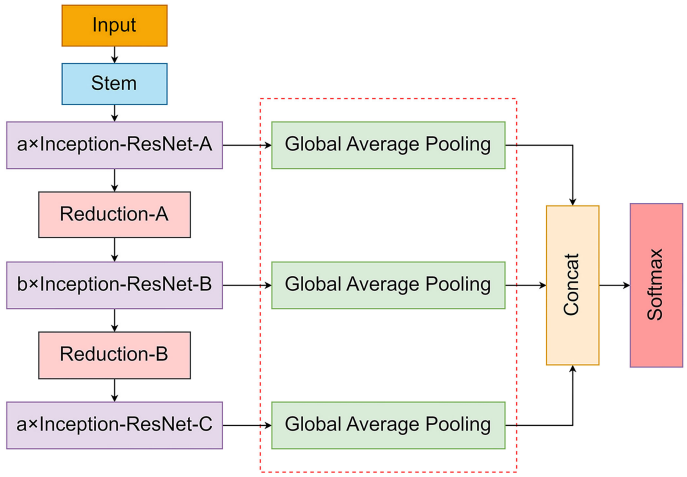


FIGURE 8
Network structure with added feature fusion.

3.2 Comparison of machine learning experiments

To verify the performance of the improved model, three features of maize seed (shape, color, and texture) were extracted using traditional machine vision techniques, and the three types were combined to quantify the images more effectively. The logistic regression, k-nearest neighbor, decision tree, random forest, gaussian naïve bayes, and support vector machine models were used for classification, and the common lightweight convolutional neural networks DenseNet, MobileNet, ShuffleNet, and Inception, and the original Inception-ResNet were selected for comparison. The dataset was divided into training and test sets in a 4:1 ratio, and the classification accuracies of the different models on the test set are shown in Figure 11A. The highest accuracy with the test set occurred with the improved Inception-ResNet model, with 99.49%, which was an increase of 2.03 percentage points over the accuracy of the original model of 97.46% and an increase of 5.07 percentage points compared to the LR model, which had the highest machine-learning recognition accuracy for the test set of 94.42%. In addition, a confusion matrix was chosen as the visual presentation tool to evaluate the quality of the classification models, and the matrices for several models of the test set are shown in Figure 11B. The plots showed that the original Inception-ResNet model had an accuracy of 96.67% for the identification of seeds with a defective appearance and 98.13% for the identification of seeds with a good appearance. The confusion matrix for the improved Inception-Resnet model showed the best results, with a significant improvement and a classification accuracy of the test set seeds of up to 100%; therefore, this model provided an accurate identification of seed appearance quality.

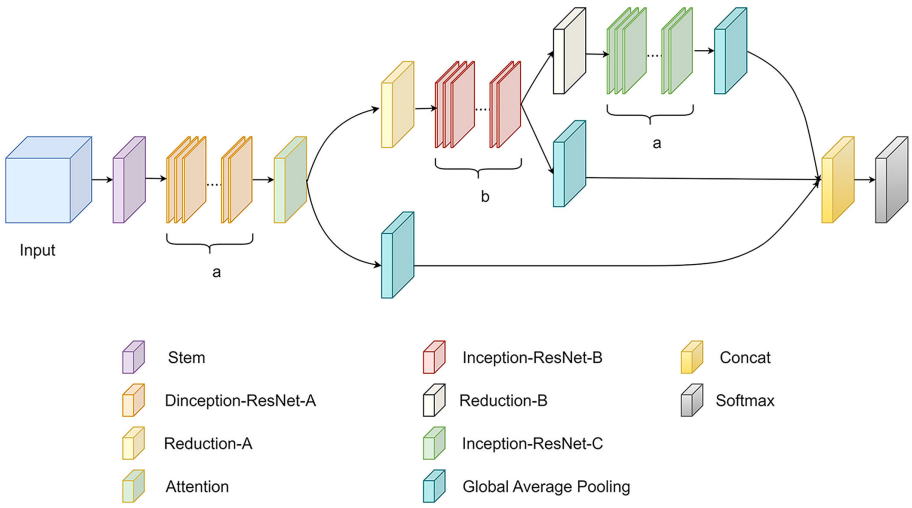


FIGURE 9
Improved model structure.

TABLE 3 Comparison of improving module performance.

	Model	Average Accuracy/%	Average Precision/%	Average Recall/%	F1/%	Detection time/s
1	Inception-Resnet	88.10	90.26	88.10	89.17	2.59
2	Inception-Resnet (+depthwise separable convolution)	89.68	91.33	89.67	90.49	2.56
3	Inception-Resnet (+ECANet)	95.96	92.46	91.26	91.86	2.68
4	Inception-Resnet (+feature fusion)	93.65	94.28	93.65	93.96	2.62
5	Inception-Resnet (+depthwise separable convolution + ECANet)	92.06	93.05	92.06	92.55	2.77
6	Improved Inception-ResNet	96.03	96.27	96.03	96.15	2.44

3.3 Comprehensive testing performance

To evaluate the effectiveness of the improved Inception-ResNet in detecting seed appearance quality, ten sets of image data with good and bad co-existing seeds were selected for the final validation and performance of the six detection algorithms, DenseNet, MobileNet, ShuffleNet, Inception, Inception-ResNet, and improved Inception-ResNet. The models were evaluated using the criteria of average accuracy, precision, recall and the reconciled average evaluation F1 value. The results are shown in Table 4, and the highest average accuracy of Inception-ResNet after improvement reached 96.03%, which was 3.23 – 11.11 percentage points higher than those of the other models, the average precision reached 96.27%, an improvement of 3.46 – 8.00 percentage points compared to other networks. Similarly, the average recall was 3.22 – 11.11 percentage points greater than the other models, with 96.03%, while the reconciled average evaluation F1 value reached 96.15%, which was 3.34 – 9.59 percentage points higher than the other algorithms and the average time to achieve real-time detection was 2.44 seconds per detected image.

3.4 Model performance

Combined with the watershed algorithm to obtain the specific location coordinates of the corn seeds for defect detection, the detection results of several models are shown in Figure 12. The improved Inception-ResNet model had a significantly greater number of cases with target confidence levels of 1.0, with approximately 88% of the seed identifications showing a level approaching 1.0, which is a six-percentage point improvement over the level of the original model. A comparison of the detection results is shown in Table 5, which shows that the improved Inception-ResNet model had the greatest number of correct detections and highest overall recognition accuracy, thereby allowing for an accurate detection of the appearance quality of corn seeds to achieve the desired results. In the actual detection process, there are also cases of misdetection, such as the possibility of detecting bad seed as good seed when the area of damage is small. When the seeds are randomly placed at an inappropriate Angle or there is serious adhesion leading to area coverage between seeds, it is possible to detect good seeds as bad seeds. This also indicates that there is room for model refinement and we need to follow up to explore more accurate detection models.

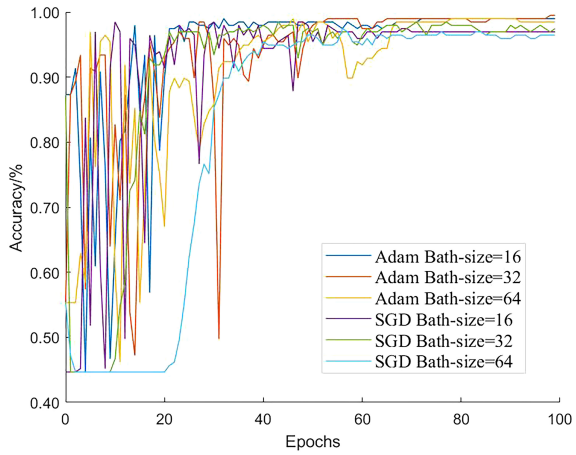


FIGURE 10
Comparison of parameters.

10.3389/fpls.2023.1249989

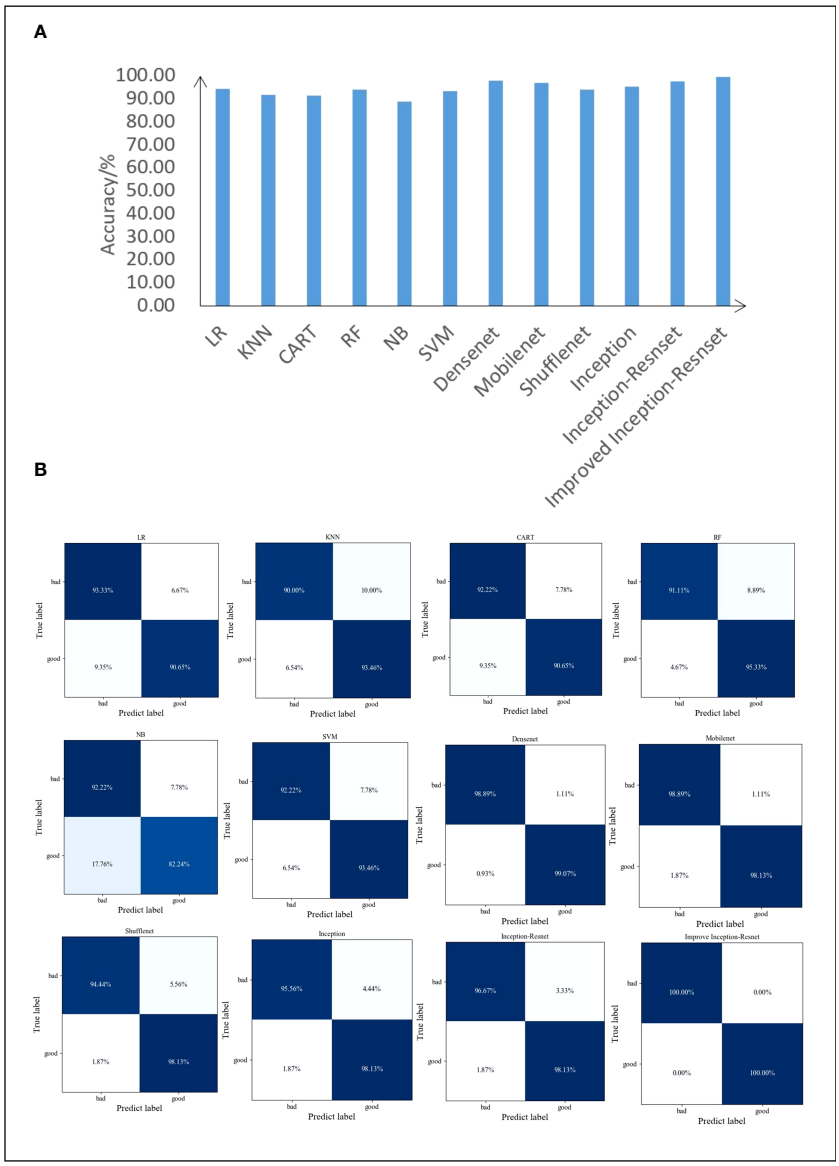


FIGURE 11
Test result **(A)** Comparison of accuracy between different model test sets and **(B)** Confusion matrix of different models.

4 Discussion

We proposed an improved model based on Inception-Resnet for corn seed appearance quality detection, which can realize the purpose of accurate and real-time detection. However, the model is domain dependent and mainly focuses on the detection of seed appearance quality, which makes it difficult to directly migrate the network to other domains. This is mainly due to the fact that the network structure modifications are performed based on seed features, which are difficult to guarantee that they will respond well on datasets from other domains. With the development of agricultural digitization, it can be applied to seed intelligent sorting equipment to guarantee the quality of seeds. We can do further research to solve some problems, firstly, the data collection was conducted indoors, in the future, we can try to train the model directly for the dataset collected in the actual outdoor environment to improve the generalization ability of the model; secondly, we can design a lighter network structure, which can ensure the stability of the algorithm and efficiently deploy it on different platforms while pursuing high performance; finally, we only study the appearance damage of corn seeds, and in the future we can continue to explore more specific defects on the seed surface, such as diseases and pests, and further promote the development of the seed industry.

5 Conclusions

To improve the accuracy of seed appearance quality assessment, this study proposed an improved Inception-ResNet model based on the Inception-ResNet algorithm for identifying appearance defects

10.3389/fpls.2023.1249989

TABLE 4 Comparison of detection model performances.

Model	Average Accuracy/%	Average Precision/%	Average Recall/%	F1/%	Detection time/s
DenseNet	84.92	88.27	84.92	86.56	2.86
MobileNet	87.30	89.75	87.30	88.51	1.81
ShuffleNet	88.87	90.24	88.88	89.55	1.91
Inception	92.80	92.81	92.81	92.81	2.55
Inception-ResNet	88.10	90.26	88.10	89.17	2.59
Improved Inception-ResNet	96.03	96.27	96.03	96.15	2.44

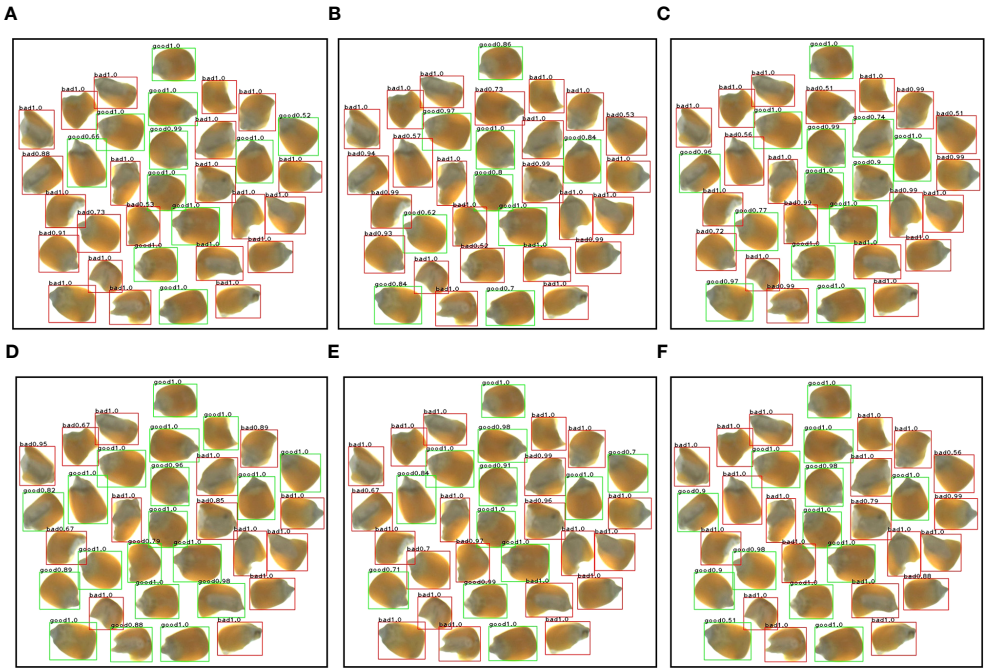


FIGURE 12
Comparison of detection results of multiple algorithms **(A)** DenseNet **(B)** MobileNet **(C)** ShuffleNet **(D)** Inception **(E)** Inception-ResNet **(F)** Improved Inception-ResNet.

of corn seeds and obtained the following conclusions: (1) Taking advantage of the small number of depthwise separable convolutional parameters as opposed to the standard convolution in the Inception-ResNet module reduced the large number of parameters generated by overlaying the Inception-ResNet module and the requirement for hardware resources. (2) Introducing the Efficient Channel Attention Network strengthened the ability to learn key information and avoided the problem of excessive information storage and information overload in the model. Simultaneously, the output feature maps were fused to obtain richer feature information to enhance the network generalization ability and improve network performance. (3) The detection effect of the method proposed in this study was superior to the other models tested, with an average accuracy of 96.03%, average precision of 96.27%, average recall of 96.03%, F1 value of 96.15%, and detection speed for a single corn seed image of approximately 2.44 seconds. The performance index improved significantly with high performance stability, providing a theoretical basis for subsequent seed quality detection.

TABLE 5 Comparison of test results.

Model	Correct number	Miss number
DenseNet	27	6
MobileNet	25	8
ShuffleNet	25	8
Inception	27	6
Inception-ResNet	28	5
Improved Inception-ResNet	29	4

Data availability statement

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

Author contributions

CS, BP, and XF conceived the idea, proposed the method. HW and YZ set up the platform and collected data. CS and BP write the code, test the results and write the manuscript. LS and XS contributed to the validation results. XF and XS provided guidance for the writing of this paper. All authors contributed to the article and approved the submitted version.

Funding

This study was supported by the National Natural Science Foundation of China (32072572), the earmarked fund for CARS (CARS-23), the Talent Recruiting Program of Hebei Agricultural University (YJ2021056), and the research project on Basic Research Business Expenses of Provincial Higher Education Institutions in Hebei Province (KY2022019).

Acknowledgments

Thanks to the teachers in the project room for their guidance and the help of the students. And we would like to thank Editage (www.editage.cn) for English language editing.

Conflict of interest

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

Publisher's note

References

Ajit, A., Acharya, K., and Samanta, A. (2020). *A review of convolutional neural networks[C]//2020 international conference on emerging trends in information technology and engineering (ic-ETITE)* (Vellore, India: IEEE 2020 International Conference on Emerging Trends in Information Technology and Engineering (ic-ETITE)). doi: 10.1109/ic-ETITE47903.2020.049

Alexandros, O., Cagatay, C., and Ayalew, K. (2023). Deep learning for crop yield prediction: a systematic literature review. *New Z. J. Crop Hortic. Sci.* 51 (1), 1–26. doi: 10.1080/01140671.2022.2032213

Ali, A., Qadri, S., Mashwani, W. K., Brahim, B., Naeem, S., Rafique, S., et al. (2020). Machine learning approach for the classification of corn seed using hybrid features. *Int. J. Food Properties* 23 (1), 1097–1111. doi: 10.7779/jksnt.2020.40.1.47

Anikó, N., and Miklós, N. (2022). Crop yield prediction in precision agriculture. *Agronomy* 12 (10), 2460–2460. doi: 10.3390/agronomy13071812

Ansari, N., Ratri, S. S., Jahan, A., Ashik-E-Rabbani, M., and Rahman, A. (2021). Inspection of paddy seed varietal purity using machine vision and multivariate analysis. *J. Agric. Food Res.* 3, 100109. doi: 10.1016/j.jafr.2021.100109

Cai, C., Wang, Q., Cai, W., Yang, Y. X., Hu, Y. H., Li, L. J., et al. (2023). Identification of grape leaf diseases based on VN-BWT and Siamese DWOAM-DRNet. *Eng. Appl. Artif. Intell.* 123 (PB). doi: 10.1016/J.ENGAPPAI.2023.106341

Cong, S., and Zhou, Y. (2022). A review of convolutional neural network architectures and their optimizations. *Artif. Intell. Rev.* 56 (3), 1905–1969. doi: 10.3390/app10103443

Chen, Y. H., An, X. S., Gao, S. M., Li, S. J., and Kang, H. W. (2021). A deep learning-based vision system combining detection and tracking for fast on-line citrus sorting. *Front. Plant Sci.* 12. doi: 10.3389/fpls.2021.622062

Fang, S., Wang, Y., Zhou, G., Chen, A., Cai, W., Wang, Q., et al. (2022). Multi-channel feature fusion networks with hard coordinate attention mechanism for maize disease identification under complex backgrounds. *Comput. Electron. Agric.* 203, 107486. doi: 10.1016/j.compag.2022.107486

Farooqui, N. A., and Ritika, R. (2019). An identification and detection process for leaves disease of wheat using advance machine learning techniques. *Bioscience Biotechnol. Res. Commun.* 12 (4), 1081–1091. doi: 10.21786/bbrc/12.4/31

Feng, X., Zhao, C., Wang, C., Wu, H., Miao, Y., and Zhang, J. (2000). *A vegetable leaf disease identification model based on image-text cross-modal feature fusion frontiers in plant science*, (Laussane, Switzerland: Frontiers in Plant Science) 13, 918940. doi: 10.3389/fpls.2022.918940

Gao, X. H., and Liu, B. (2016). Design and experiment of fresh corn quality detection classifier based on machine vision. *Trans. Chin. Soc. Agric. Eng.* 32 (01), 298–306. doi: 10.11975/j.issn.1002-6819.2016.01.041

Gao, F. F., Fu, L. S., Zhang, X., Majeed, Y., Li, R., Karkee, M., et al. (2020). Multi-class fruit-on-plant detection for apple in SNAP system using Faster R-CNN. *Comput. Electron. Agric.* 176. doi: 10.3389/fpls.2021.622062

Hou, J. M., Yao, E. C., and Zhu, H. J. (2020). Classification of castor seed damage based on convolutional neural network. *Trans. Chin. Soc. Agric.* 51 (S1), 440–449. doi: 10.6041 /j.issn.1000-1298.2020.S1.052

Huang, S., Fan, X. F., Sun, L., Shen, Y. L., and Suo, X. S. (2019). Research on classification method of maize seed defect based on machine vision. *J. Sensors* 2019, 1–9. doi: 10.1155/2019/2716975

Kamilaris, A., and Prenafeta-Boldu, F. X. (2018). A review of the use of convolutional neural networks in agriculture. *J. Agric. Sci.* 156 (3), 312–322. doi: 10.1016/j.compag.2022.107605

Kavita, J., Pratistha, M., Sanchit, J., and Sukriti, N. (2023). Crop yield prediction using machine learning and deep learning techniques. *Proc. Comput. Sci.* 218, 406-417. doi: 10.1016/J.PROCS.2023.01.023

Kim, D. H., Boo, S. B., Hong, H. C., Yeo, W. G., and Lee, N. Y. (2020). Machine vision-based defect detection using deep learning algorithm. *J. Korean Soc. nondestructive testing* 40 (1), 47–52. doi: 10.7779/jksnt.2020.40.1.47

Laabassi, K., Belarbi, M. A., Mahmoudi, S., Ahmed, M. S., Ferhat, K., et al. (2021). Wheat varieties identification based on a Deep Learning approach. *J. Saudi Soc. Agric. Sci* 20 (5), 281-289. doi: 10.1016/j.jssas.2021.02.008

Li, L., Hu, D. Y., Tang, T. Y., and Tang, Y. L. (2022). Nondestructive testing of the quality of different grades of creamy strawberries based on Laida algorithm. *J. Food Process. Preservation* 46 (11). doi: 10.1111/JFPP.17008

Liang, W. J., Zhang, H., Zhang, G. F., and Gao, H. X. (2019). Rice blast disease recognition using a deep convolutional neural network. *Sci. Rep.* 9 (1), 1–10. doi: 10.1038/s41598-019-38966-0

Liu, S., Qiao, Y. L., Li, J. W., Zhang, H. T., Zhang, M. K., and Wang, M. L. (2022). An improved lightweight network for real-time detection of apple leaf diseases in natural scenes. *Agronomy* 12 (10). doi: 10.3390/AGRONOMY12102363

Liu, J., Wang, X., Miao, W., and Liu, G. X. (2022). Tomato pest recognition algorithm based on improved YOLOv4. *Front. Plant Sci.* 13, 814681. doi: 10.1109/ACCESS.2019.2914929

Liu, M. Y., Zhou, G. X., Chen, A. B., Li, L. J., and Hu, Y. H. (2023). Identification of tomato leaf diseases based on LMBRNet. *Eng. Appl. Artif. Intell.* 123 (PA).

Lu, C. Q., Li, Z. Y., Shi, S. Y., and Peng, S. L. (2022). Research on target defect detection based on machine vision. *J. Physics: Conf. Ser.* 2219 (1). doi: 10.1088/1742-6596/2219/1/012060

Ma, J. J., Wang, K. Q., Zheng, Y. X., Cai, K., and Lin, Q. Y. (2021). Research on the appearance quality detection and classification of peanut seeds based on machine Vision. *Anhui Agric. Sci.* 49 (10), 225–227, 231. doi: 10.11975/j.issn.1002-6819.2022.14.033

Mamalet, F., and Garcia, C. (2012). *Simplifying convnets for fast learning[C]// International Conference on Artificial Neural Networks* (Berlin, Heidelberg: Springer), 58–65.

Naranjo-Torres, J., Mora, M., Hernández-García, R., Barrientos, R. J., Fredes, C., and Valenzuela, A. (2020). A review of convolutional neural network applied to fruit image processing. *Applied Sciences* 10 (10), 3443. doi: 10.3390/app10103443

Peng, J., Chen, Y. H., Liu, B., He, D. J., and Liang, C. Q. (2019). Real-Time detection of apple leaf diseases using deep learning approach based on improved convolutional neural networks. *IEEE Access* 7, 59069–59080. doi: 10.1109/ACCESS.2019.2914929

Peng, D., Li, W., Zhao, H., Zhou, G., and Cai, C. (2023). Recognition of tomato leaf diseases based on DIMPCNET. *Agronomy* 13, 1812. doi: 10.3390/agronomy13071812

Prajapati, H., Shah, J., and Dabhi, V. (2018). Detection and classification of rice plant diseases. *Intelligent Decision Technol.* 2), 1–18. doi: 10.3233/idt-180338

Rahmani, M. E., Amine, A., and Hamou, R. M. (2016). Supervised machine learning for plants identification based on images of their leaves. *Int. J. Agric. Environ. Inf. Syst.* 7 (4), 17–31. doi: 10.4018/IJAEIS.20211001.OA5

Rong, D., Xie, L. J., and Ying, Y. B. (2019). Computer vision detection of foreign objects in walnuts using deep learning. *Comput. Electron. Agric.* 162, 1001 – 1010. doi: 10.1016/j.compag.2019.05.019

Sj, A., Shma, B., Fjv, A., and Am, C. (2021). Computer-vision classification of corn seed varieties using deep convolutional neural network. *J. stored products Res.doi: 10.1016/j.jspr.2021.101800.* 92, 101800. doi: 10.1016/j.jspr.2021.101800

Sharaff, A., and Patel, B. (2021). Rice crop disease prediction using machine learning technique . *Int. J. Agric. Environ. Inf. Syst.* 12 (4), 1–15. doi: 10.4018/ IJAEIS.20211001.OA5

Tan, S. Y., Ma, X., Mai, Z. J., Qi, L., and Wang, Y. W. (2019). Segmentation and counting algorithm for touching hybrid rice grains. *Comput. Electron. Agric.* 162, 493–504. doi: 10.1016/j.compag.2019.04.030

Tiwari, V., Joshi, R. C., and Dutta, M. K. (2021). Dense convolutional neural networks based multiclass plant disease detection and classification using leaf images. *Ecol. Inf.doi* 101289. doi: 10.1016/J.ECOINF.2021.101289

Wang, L., Liu, J., Zhang, J., Wang, J., and Fan, X. (2022). Corn seed defect detection based on watershed algorithm and two-pathway convolutional neural networks. *Front. Plant Sci.* 13. doi: 10.3389/fpls.2022.730190

Wang, Z. L., Huang, W. Q., Tian, X., Long, Y., Li, L. J., and Fan, S. X. (2022). Rapid and Non-destructive Classification of New and Aged Maize Seeds Using Hyperspectral Image and Chemometric Methods. *Front. Plant Sci.* 13. doi: 10.3389/FPLS.2022.849495

Wang, Z., and Xiao, Z. (2020). *Product quality detection and recognition based on vision and deep learning[C]//2020 35th youth academic annual conference of chinese association of automation (YAC)* (China: Youth Academic Annual Conference of Chinese Association of Automation (YAC)). doi: 10.1109/yac51587.2020.9337602

Wang, X., Yang, W., Lv, Q., Huang, C., Liang, X., Chen, G., et al (2022). Field rice panicle detection and counting based on deep learning. *Front. Plant Sci.* 13, 966495. doi: 10.3389/fpls.2022.966495

Wu, J. Y., Li, B., and Wu, Z. L. (2019). Detection of crop pests and diseases based on deep convolutional neural network and improved algorithm[P]. *Mach. Learn. Technol.*, 20–27. doi: 10.1145/3340997.3341010

Yuan, H. Y., Cheng, J. P., Wu, Y. R., and Zeng, Z. Y. (2022). Low-res MobileNet: An efficient lightweight network for low-resolution image classification in resource-constrained scenarios. *Multimedia Tools Appl.doi:10.1007/S11042-022-13157-8.* 81 (27), 38513-38530. doi: 10.1007/S11042-022-13157-8

Zhang, C., Zhang, Y. Y., Yan, T. Y., Bai, X. L., Xiao, Q. L., Gao, P., et al. (2020). Application of near-infrared hyperspectral imaging for variety identification of coated maize kernels with deep learning. *Infrared Phys. Technol.* 111 (20), 103550. doi: 10.1016/j.compag.2019.04.030

Zhao, Y. Q., Yang, H. L., Zhang, Y., Yang, Y. K., Yang, Y., and Sai, M. (2022). Detection of defective Arabica green coffee beans based on feature combination and SVM. *Trans. Chin. Soc. Agric. Eng. (Transactions CSAE)* 38 (14), 295–302. doi: 10.11975/j.issn.1002-6819.2022.14.033

frontiers | Frontiers in Plant Science

TYPE Original Research
PUBLISHED 29 August 2023
DOI 10.3389/fpls.2023.1248598

OPEN ACCESS

EDITED BY
Jianwei Qin,
Agricultural Research Service (USDA), United States

REVIEWED BY
Pappu Kumar Yadav,
University of Florida, United States
Ebenezer Olaniyi,
Mississippi State University, United States
Princess Tiffany D. Mendoza,
Kansas State University, United States

*CORRESPONDENCE
Chunjiang Zhao
Zhaocj@nercita.org.cn
Xi Tian
tianx2019@sina.com

RECEIVED 27 June 2023
ACCEPTED 11 August 2023
PUBLISHED 29 August 2023

CITATION
Fan Y, An T, Wang Q, Yang G, Huang W, Wang Z, Zhao C and Tian X (2023) Non-destructive detection of single-seed viability in maize using hyperspectral imaging technology and multi-scale 3D convolutional neural network. *Front. Plant Sci.* 14:1248598. doi: 10.3389/fpls.2023.1248598

Non-destructive detection of single-seed viability in maize using hyperspectral imaging technology and multi-scale 3D convolutional neural network

Yaoyao Fan[1,2], Ting An[2], Qingyan Wang[2], Guang Yang[2], Wenqian Huang[2], Zheli Wang[2], Chunjiang Zhao[1,2]* and Xi Tian[2]*

[1]College of Information and Electrical Engineering, Shenyang Agricultural University, Shenyang, China, [2]Intelligent Equipment Research Center, Beijing Academy of Agriculture and Forestry Sciences, Beijing, China

The viability of Zea mays seed plays a critical role in determining the yield of corn. Therefore, developing a fast and non-destructive method is essential for rapid and large-scale seed viability detection and is of great significance for agriculture, breeding, and germplasm preservation. In this study, hyperspectral imaging (HSI) technology was used to obtain images and spectral information of maize seeds with different aging stages. To reduce data input and improve model detection speed while obtaining more stable prediction results, successive projections algorithm (SPA) was used to extract key wavelengths that characterize seed viability, then key wavelength images of maize seed were divided into small blocks with 5 pixels ×5 pixels and fed into a multi-scale 3D convolutional neural network (3DCNN) for further optimizing the discrimination possibility of single-seed viability. The final discriminant result of single-seed viability was determined by comprehensively evaluating the result of all small blocks belonging to the same seed with the voting algorithm. The results showed that the multi-scale 3DCNN model achieved an accuracy of 90.67% for the discrimination of single-seed viability on the test set. Furthermore, an effort to reduce labor and avoid the misclassification caused by human subjective factors, a YOLOv7 model and a Mask R-CNN model were constructed respectively for germination judgment and bud length detection in this study, the result showed that mean average precision (mAP) of YOLOv7 model could reach 99.7%, and the determination coefficient of Mask R-CNN model was 0.98. Overall, this study provided a feasible solution for detecting maize seed viability using HSI technology and multi-scale 3DCNN, which was crucial for large-scale screening of viable seeds. This study provided theoretical support for improving planting quality and crop yield.

KEYWORDS

viability detection, maize seeds, hyperspectral imaging, YOLOv7 model, 3D convolution neural network

1 Introduction

Single-seed sowing is a crucial strategy to boost corn production, save seeds, and reduce labor, but it demands high-quality seeds (Li et al., 2017). On October 11th, 2020, a new standard has been released by China, which raises the germination rate index for single-seed sowing from 85% to 93%. The viability is a critical indicator for evaluating the quality and practicality of seed. Assessment of seed viability could ensure each seed has the potential for germination and healthy growth and promotes the popularization of single-seed sowing. This not only facilitates mechanized sowing and reduces the laboriousness of manual interplanting and seedling transplantation, but also significantly reduces the amount of seed used and conserves a considerable amount of seed production area (Liang et al., 2020). Therefore, the determination of seed viability is of utmost importance in reducing the cost and time loss resulting from planting failures and conserving human resources.

Seed viability is a quality characteristic at the individual level rather than a quantitative trait at the population level. Loss of viability among individuals in the same population is not synchronous, making it challenging to detect the viability of single-seed. According to the International Seed Testing Association (ISTA) rules (Association, I.S.T, 1999), common methods for seed viability detection include germination and staining (Cheng et al., 2023). The conventional germination method is the most accurate, but it is time-consuming and requires a lot of material resources. On the other hand, staining is only suitable for a small number of samples. Therefore, it is necessary to develop a rapid-nondestructive technique for single-seed viability detection in large quantities.

In the field of seed quality detection, hyperspectral imaging technology has been widely utilized. However, research on seed viability detection is relatively limited. Jannat Yasmin et al. (2022) presented an online detection system of watermelon seed viability based on longwave near-infrared (LWNIR) HSI, demonstrating its potential application in predicting seed viability. Wang et al. (2021) developed the discrimination models of seed viability using the feature wavelengths and full wavelengths of the visible and shortwave near-infrared (Vis-SWNIR), the result revealed that both models attained an accuracy rate surpassing 95%, suggesting that the seeds with different aging stages exhibited unique spectral features, and the characteristic wavelengths can effectively provide the key information of Zea mays seed quality. Pang et al. (2021) conducted a germination experiment on maize seeds with different aging stages, a 2D convolutional neural network (2DCNN) model was developed by combing deep learning algorithms with hyperspectral technology. The accuracy of this model reached 99.96%, which was significantly higher than machine learning and one-dimensional convolutional neural network (CNN). It was worth pointing out that the model demonstrated a relatively fast convergence speed, which highlighted the feasibility and effectiveness of combining deep learning with hyperspectral technology to determine the viability of single-seed. Ambrose et al. (2016) investigated the feasibility of using HSI technology to differentiate the viability of maize seeds. One group of maize samples was subjected to microwave heat treatment, while the other group served as the control. PLS-DA was employed to classify the heat-treated (aged) and untreated (normal) maize seeds. The results showed that the classification model achieved the highest classification accuracy in the LWNIR region, with calibration set accuracy of 97.6% and prediction set accuracy of 95.6%. These studies achieve high accuracy by predicting the aging level or treatment condition of seeds instead of the actual results of germination experiments. And they mainly rely on overall image information for seed viability classification. However, they overlook the significance of local information within seeds and fail to consider subtle variations and characteristics in different seed regions.

Generally, the evaluation of germination rate of seeds mainly depends on manual labor, which is time-consuming and cumbersome. Zhao et al. (2022) proposed a detection method for the germination rate of rice seeds using deep learning models, which took an average of 0.011 seconds for each image while achieving a mAP of 0.9539, meeting the demands of real-time detection, indicating that the YOLO-r model had great potential for rapidly and precisely determining the germination status of seeds. Bai et al. (2023) developed an improved discriminative approach for the detection of seed germination using YOLOv5. This technique enables the swift evaluation of parameters such as wheat seed germination rate, germination potential, germination index, and average germination days.

The emergence ability of seedlings is crucial for seed growth and crop yield improvement (Cui et al., 2020). In recent studies, significant progress has been made in correlating seed germination ability and seedling growth through various measurement methods. However, traditional manual measurement techniques for assessing parameters like bud length have been found to be inefficient and prone to errors due to the complex and twisted nature of buds. To address this issue, Adegbuyi and Burris (1988) found there was a significant correlation between seed germination ability and seedling growth by measuring comprehensive growth parameters. However, manual measurement method of bud length is inefficient and error-prone due to their curved and twisted nature. Gaikwad et al. (2019) developed a semi-automated tool for measuring leaf length, width, and area. Abdelaziz Triki et al. (2021) used the Mask R-CNN algorithm to effectively segment and measure leaf characteristics and obtained an error rate of around 5%. An enhanced algorithm based on the mask RCNN was introduced by Shen et al. (2023) to recognize defective wheat kernels. The experimental outcomes showed that this refined algorithm facilitated quicker and more precise detection of unsound kernels, effectively tackling issues linked to kernel adhesion. Masood et al. (2021) propose an automated method that utilizes the Mask RCNN model to achieve precise localization and segmentation of brain tumors. Cui et al. (2022) constructed a recognition model using hyperspectral data and feature extraction algorithms to predict maize root length, showing a significant correlation between root length and viability. Therefore, it is of great significance to measure and predict the seed viability using computer technology.

The above study highlighted the significance of seed viability determination and emphasized the need of developing rapid and

non-destructive technology for single-seed viability detection. HSI has been established as a useful tool for seed quality detection, and the integration of deep learning and hyperspectral technology can establish an effective seed viability detection model. However, previous studies commonly used relatively simple models, and lacking the prediction model of maize seeds viability developed by 3DCNN and hyperspectral images. This study proposed an improved method for identifying the viability of maize seeds based on germination experiments. The aim of the study is to explore the potential of using hyperspectral images and 3DCNN to identify the viability of maize seeds. Specifically, the objectives are to: (1) select characteristic wavelengths that represent seed viability, (2) combine HSI with 3DCNN to establish the optimal classification model for maize seed viability, (3) evaluate the feasibility of using YOLOv7 model instead of the human eye to determine the seed germination status, (4) evaluate the ability of Mask R-CNN in bud segmentation and bud length prediction.

2 Materials and methods

2.1 Maize sample preparation

2.1.1 Aging experiment

Due to the high quality and the resistance to multiple stressors, "Jingke 968" maize is extensively cultivated in eastern and northern China. Therefore, it was selected as the experiment sample in this study. To ensure the accuracy of the experiment, seeds with uniform size and shape were manually selected, then all seeds were disinfected by soaking them in a 0.5% sodium hypochlorite solution for 5 minutes, followed by rinsing with distilled water five times, and air-dried under natural conditions.

To simulate the natural aging process of seeds, the experiment samples were artificially aged. All seeds were exposed to high temperature and high humidity conditions (45 °C and a relative humidity of 95%) and stirred twice a day to ensure uniform exposure (Zhang et al., 2020). 150 maize seeds were taken out randomly at aging 2, 4, 6, and 8 days, respectively. Additionally, 150 untreated seeds were selected as the control group (CK). Therefore, a total of 750 maize seeds within five aging stages were obtained and used for subsequent experimentation.

2.1.2 Hyperspectral imaging system

Two HSI systems, the Vis-SWNIR and LWNIR, have been built in the Intelligent Detection Laboratory of the China Agricultural Equipment Technology Research Center (Fan et al., 2018). The Vis-SWNIR system is capable of acquiring hyperspectral images within the wavelength range of 327-1098 nm, encompassing 1000 spectral variables, while the LWNIR system can capture images within the range of 930-2548 nm, containing 256 spectral variables. The Vis-SWNIR system includes an imaging spectrometer, an electron-multiplying charge-coupled device camera with a resolution of 502×500, a camera lens, and a spectraCube data acquisition software. Similarly, the LWNIR system includes an imaging spectrometer, a charge-coupled device camera with a resolution of 320×256, a camera lens, and a spectral acquisition software (Tian et al., 2021). And the acquisition software of both systems was developed using LabVIEW (National Instruments Inc., Austin, TX, USA) to facilitate the acquisition of spectral images, as well as to manage the camera and motor operations. Both systems share two 300-watt halogen lamps to provide stable illumination. In addition, an electrically operated moving platform and a computer are available for sample placement (Capable of accommodating up to 96 samples simultaneously) and hyperspectral image acquisition (Figure 1A) (Liu et al., 2022).

To ensure the accuracy and reliability of the hyperspectral images (E_{raw}), calibration operation is essential to eliminate the effects of uneven illumination of the light source and camera dark current changes (An et al., 2022). The calibration operation involved using a white reflection board (with a reflectance of 99%) (E_w) to acquire a standard white reference image in the same sampling environment as the sample, while turning off the light source and covering the lens to obtain a black reference image (with a reflectance of 0%) (E_d). The calibrated image can be calculated using the following formula:

$$E_c = \frac{E_{raw} - E_d}{Ew - Ed} \quad (1)$$

After calibration, in the Vis-SWNIR region, a subset of 347 spectral variables within the 420-1000 nm range was selected for further analysis, considering the abundance of spectral data and the presence of duplicate information in adjacent spectra. On the other hand, in the near-infrared region, due to the limited number of available bands, all spectral variables (256) were directly included in the analysis. To separate maize seeds from the background, a mask was applied to segment the hyperspectral image. The gray-scale images at 801 nm and 1098 nm were selected as the mask images for the Vis-SWNIR and LWNIR bands, respectively. The average spectral curves were obtained by calculating the mean reflectance under the mask. Lastly, in order to eliminate the influence of the instrument, the Savitzky-Golay (SG) and Standard Normal Variate (SNV) methods were utilized to preprocess the spectra.

2.1.3 Standard germination test

A transparent box measuring 25cm×25cm was used as a germination chamber, and 75 seeds were placed in each box. A total of 10 boxes were used in the experiment. Prior to the germination test, the germination boxes were sterilized with 75% ethanol (Suksungworn et al., 2021), and three layers of gauze were placed in each germination box to provide continuous moisture for the seeds. A black gauze was placed on the top layers as the background for photography (Figure 2A). An equal amount of distilled water was added to each box, and the temperature was set to 25°C with 12-hour intervals of light and dark (Figure 2B). Throughout the 7-day germination experiment (Long et al., 2022), the germination progress of maize seeds was monitored daily at specific time intervals. According to the ISTA standard, the germination rate was determined (Wang et al., 2022c).

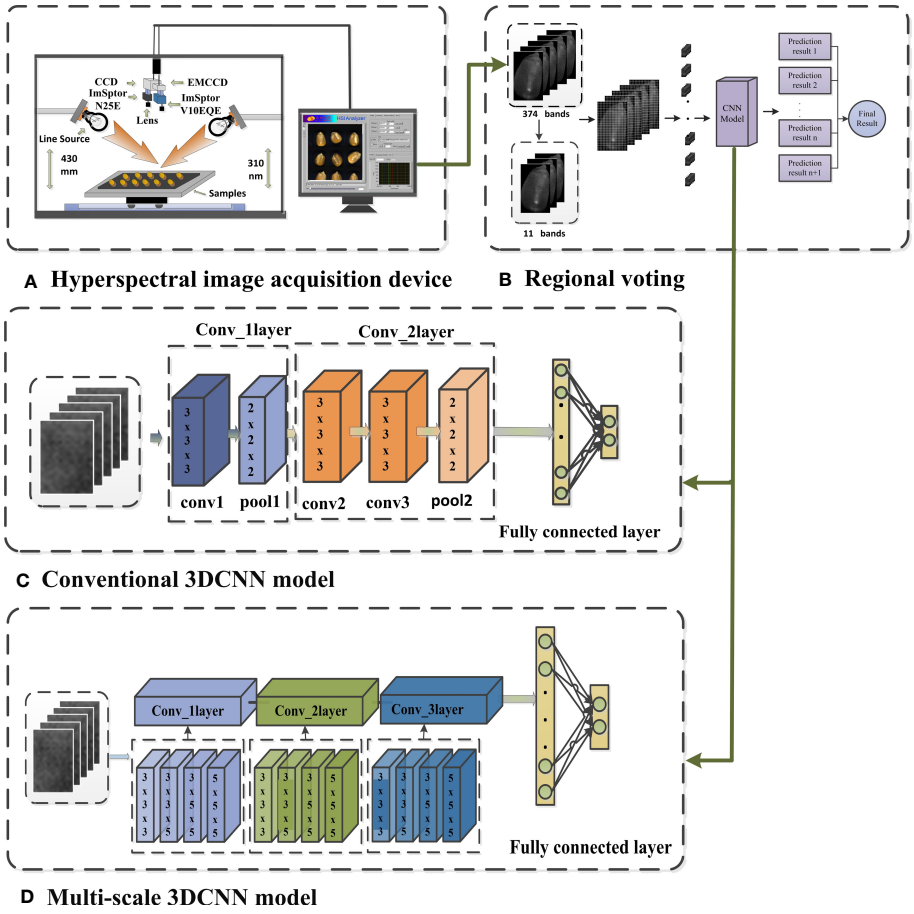


FIGURE 1
Diagram of the 3DCNN for hyperspectral image classification **(A)** Hyperspectral image acquisition device, **(B)** Regional voting, **(C)** Conventional 3DCNN model, **(D)** Multi-scale 3DCNN model.

2.1.4 RGB image acquisition

RGB images of maize seeds were captured using BASLER industrial cameras (acA1920-25um/uc, BASLER AG, Germany, 2.4 MP,100 fps) during germination test (Figure 2D) (Shen et al., 2023). An adjustable camera platform was built to ensure consistency of the images and prevent camera shake. The position of the germination box relative to the lens was kept fixed during each image capture. Indoor lighting was turned on and curtains were drawn for each capture. After placing the seeds into the boxes (Day 0), images of each box were immediately captured. Subsequently, images were captured every 15 hours for 7 consecutive days (Figure 2C). The dataset used in this study consisted of a total of 3000 maize seeds (All the captured RGB images collectively contain 3000 seeds). Among them, 2250 seeds were designated as training samples, while the remaining seeds were allocated to the test set.

2.2 Data processing

2.2.1 Successive projections algorithm

Hyperspectral data typically consists of numerous bands, and certain bands may exhibit high correlation or contain redundant information (Han et al., 2022).When training 3DCNN with full-band data, it will lead to a significant increase in the number of networks training parameters, resulting in a more complex model. This phenomenon is commonly referred to as the curse of dimensionality. (Köppen, 2000). However, band selection (Sun and Du, 2019) allows retaining spectral bands that are closely related to seed vigor assessment while removing irrelevant bands, thereby enhancing the feature extraction and discriminative capabilities of the model.

Additionally, the use of dimensionality reduction data sets can effectively reduce the complexity of the model, mitigating the risk of overfitting and enhancing the model's generalization ability and stability (Aloupogianni et al., 2023). Moreover, fewer computing resources are required during model training and inference, leading to a significant improvement in the computational efficiency of the model (XingJia et al., 2022).

Successive projections algorithm is a classical band selection method that can map high-dimensional spectral data to a low-dimensional space through multiple projections(de Almeida et al., 2018). SPA is a forward iterative search method used for selecting spectral information with minimal redundancy to address collinearity issues. The steps of SPA are shown in Table 1. The SPA is widely used in hyperspectral image processing attributed to

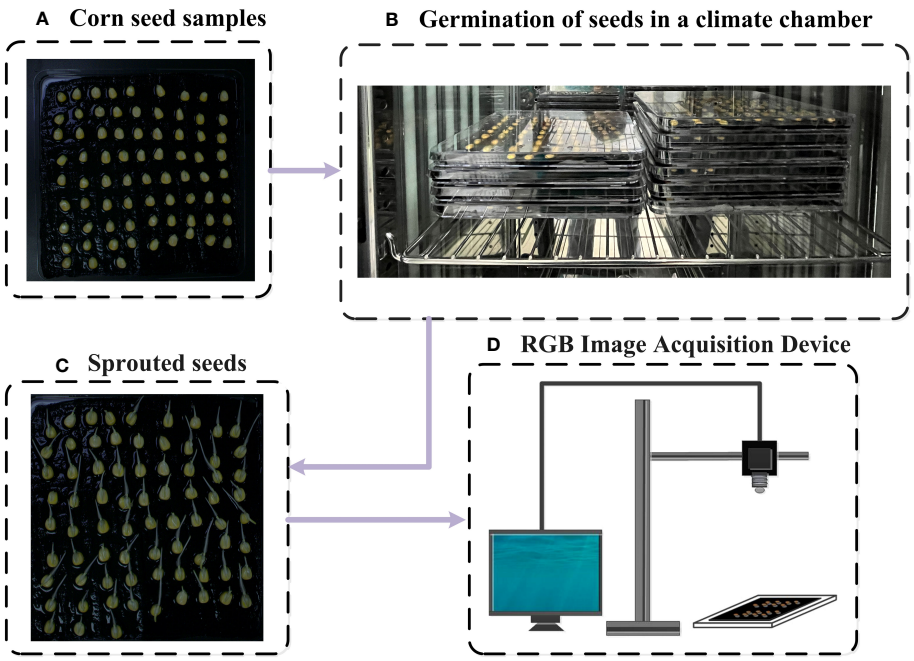


FIGURE 2
Diagram of the standard germination experiment **(A)** Corn seed samples, **(B)** Germination of seeds in a climate chamber, **(C)** Sprouted seeds, **(D)** RGB iamge acquisition device.

its advantages of fast computation speed and easy implementation (Chen et al., 2023b). Therefore, the SPA was used in this study to perform feature selection on the processed average spectra of Vis-SWNIR and LWNIR, in order to perform dimensionality reduction on the hyperspectral data.

2.2.2 Machine learning

Support Vector Machine (SVM) (Cortes and Vapnik, 1995) is a powerful algorithm for classification and regression that finds an optimal hyperplane to separate data points of different classes. It handles high-dimensional datasets, avoids overfitting, and can handle non-linear problems using kernel functions. K-Nearest Neighbor (KNN) (Zhang, 2022) is a basic algorithm that selects the K nearest samples based on their feature values and uses their labels as predictions. Subspace Discriminant Analysis (SDA) (Zhao and Phillips, 1999) is a pattern classification method that aims to find a low-dimensional subspace to maximize the separation between different classes. In this study, the aforementioned machine learning methods were used to classify the viability of maize seeds at different aging stages for optimal classification accuracy.

TABLE 1 Successive projections algorithm.

Input:Dataset with features and target variable Output:Feature subset for analysis
Step 1: For each feature in the dataset: a. Compute projection coefficients with respect to the target variable. b. Store the computed coefficients.
Step 2: Initialize an empty feature subset. **Step 3:** Repeat until desired subset size is reached or stopping criterion met: a. Find the feature with the maximum projection coefficient. b. Add the selected feature to the feature subset. c. Project out the influence of selected features on remaining features. d. Recalculate projection coefficients of remaining features.

2.2.3 Deep convolutional neural network

The CNN combines the concepts of convolutional filtering and neural networks by utilizing local receptive fields and weight sharing to reduce the number of network parameters and speed up model training (Ghaderizadeh et al., 2021). Compared to the widespread use of two-dimensional convolution, three-dimensional convolution is less commonly used in practice. However, HSI contain rich spectral information, and using two-dimensional convolution may make the interband correlations of HSIs underutilized (Ge et al., 2020). To address this issue, this study introduced a 3DCNN, which can thoroughly extract feature relationships across different feature channels (Figure 1C), thereby enabling it to concurrently extract integrated spectral and spatial features from hyperspectral imagery (Sun et al., 2022).

Before inputting hyperspectral images into the network, standardization is performed to ensure that the data is within the same scale and range, enabling the network to learn weights faster and converge more easily during training. Moreover, data standardization can help avoid the problems of gradient disappearance or explosion, and improve the stability and generalization ability of the network. To obtain multiple convolutional features of HSI, multi-scale convolution is employed in the same convolutional layer, which can acquire both global and local information. Four different convolution kernels of 3×3×3, 3×3×5, 3×5×5, and 5×5×5 were selected to extract feature information and fused on the channel. This method can enhance the classification accuracy of the model. As illustrated in Figure 1D, each

convolution kernel in the first convolution module has 16 filters, each kernel in the second convolution module has 32 filters, and each kernel in the third convolution module has 64 filters. The activation function in the three-dimensional convolution module uses Rectified Linear Unit (RELU) and is compressed by the pooling layer to reduce the amount of data and parameters, as well as alleviate the overfitting phenomenon. To ensure that the features extracted by different convolution kernels in the same module can be effectively connected, different parameters need to be set according to different situations, such as stride and padding. Finally, the output is produced through 1 fully connected layer and 1 output layer, and the output layer employs the SoftMax activation function.

To extract features from hyperspectral images of maize seeds at a more microscopic level and increase the amount of data, a window size of 5×5 was selected for segmentation (Figure 1B). To eliminate the influence of background on classification, small blocks containing 0-pixel points were discarded. As the size of maize seeds varies, the number of blocks obtained from different segments of maize seeds is also inconsistent. To address this issue, this study employed a majority principle labeling aggregation method, as Table 2.

In this study, the germination experiment showed that 404 viable samples and 346 nonviable samples were collected from 750 seeds. Given that the hyperspectral images were collected in a sequential manner based on the aging gradients of the seeds, it was crucial to maintain a balanced distribution of germinated and non-germinated samples in the test set. Therefore, a representative test set was carefully selected, consisting of 75 seeds, including the first seed, the 10th seed, the 20th seed, and so on. The remaining 675 seeds were allocated for the training phase. Through this meticulous approach, it was ensured that the test set encompassed samples from diverse categories, enabling an accurate evaluation of the classification model's performance.

2.2.4 Establishment of Mask R-CNN model for bud length detection

In order to measure the length of maize seed bud, the Mask R-CNN (He et al., 2017) (With resnet50_fpn as backbone) model was utilized to segment the bud from single-seed image firstly, then a skeleton extraction algorithm was applied to extract the skeleton of the bud (Figure 3A). Next, the bud length detection algorithm was used to remove the branches in the skeleton for obtaining the central skeleton image. Finally, the actual bud length was calculated by converting pixels to actual length (Figure 3B).

TABLE 2 Majority principle labeling aggregation method.

Input: Segmented maize seed blocks Output: Predicted potential for germination of maize seeds
Step 1: Initialize: a. Assign Label 1 to represent potential for germination. b. Assign Label 2 to indicate maize grain block affiliation.
Step 2: For each segmented maize seed block: a. Feed the block into the model for prediction of its potential for germination. b. Store the prediction result.
Step 3: For each maize seed: a. Retrieve predictions of multiple small blocks belonging to the same maize seed. b. Count the number of correct predictions. c. If more than half of the predictions are correct: - The predicted result of the maize seed is deemed correct.

Mask R-CNN is a deep learning model that combines object detection and instance segmentation. It extends Faster R-CNN by generating binary masks for each region of interest (ROI), achieving pixel-level segmentation. The network consists of three main components: a backbone network, a Region Proposal Network (RPN) responsible for generating candidate object regions, and two parallel branches dedicated to object detection and mask prediction. Mask R-CNN excels in instance segmentation, object detection, and keypoint detection, making significant contributions to computer vision advancements (Casado-García et al., 2019). The model employs a multi-task loss function, comprising classification loss (L_{cls}), bounding box loss (L_{bbox}), and predicted mask loss (L_{mask}), as represented by equations (2) to (5) (Cong et al., 2023).

$$L = L_{cls} + L_{Lbox} + L_{mask} \quad (2)$$

$$L_{cls} = \sum_i -\log\left[p_i p i_\star + \left[1 - p i_\star\right]\left[1 - p_i\right]\right] \quad (3)$$

$$L_{bbox} \frac{1}{N_{reg}} \sum_i p i_\star R \quad (t_i - t i_\star) \quad (4)$$

$$L_{mask} = -\frac{1}{m^2} \sum_{1 \le i,j \le m} \left[y_{ij}^{*} \log y_{ij} + \left[1 - y_{ij}^{*}\right] \log\left[1 - y_{ij}\right]\right] \quad (5)$$

L_{cls} measures the deviation between predicted and actual values for overall accuracy assessment. L_{bbox} quantifies the disparity between predicted and actual position parameters, assessing the model's accuracy in bud localization. L_{mask} evaluates the model's confidence in pixel-level classification using binary cross-entropy. Combining these components into a multi-task loss function allows for comprehensive evaluation across multiple tasks, resulting in enhanced overall performance.

The skeleton extraction algorithm is a technique used to extract the central line or skeleton of an object in a binary image (Fu et al., 2023). By progressively shrinking connected regions within the object contour, the algorithm produces a concise contour that provides valuable information for image processing tasks like recognition and matching. Various algorithms, such as Zhang-Suen, Morphological Thinning, and Medial Axis Transform, can be employed for this purpose. The Medial Axis Transform (MAT) algorithm, specifically, extracts the object's central line by iteratively dilating boundary pixels and identifying the nearest internal pixels as skeleton pixels. This process continues until the skeleton pixels stabilize, resulting in a stable and versatile representation suitable for subsequent image processing tasks. The MAT algorithm handles different object shapes and can process grayscale information within binary images. Seed germination images exhibit a wide range of shape features, such as bud length, curvature, and angle. However, traditional methods for measuring bud length rely on manual measurements, which are time-

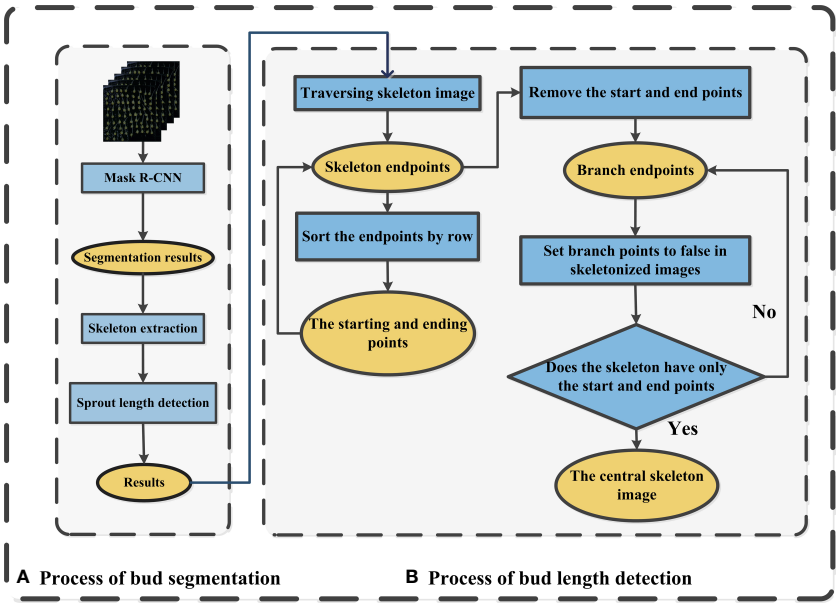


FIGURE 3
Diagram of the maize bud length detection process **(A)** Process of bud segmentation, **(B)** Process of bud length detection.

consuming and prone to significant subjective biases. The MAT (Medial Axis Transform) skeleton extraction algorithm was chosen to obtain the central line of buds. However, the resulting skeleton may contain branches that need to be eliminated to derive the center skeleton. The process of centerline skeleton extraction is illustrated in the following Figure 3B.

In this study, a transparent box with a side length of 250 mm was used as a reference to convert pixels to actual lengths in millimeters. The calculation formula is:

$$\text{Ratio} = L_{box}/1164 \quad (6)$$

Here, L_{box} represents the side length of the transparent box, and 1164 is the number of pixels corresponding to the transparent box in the image. According to the calculation formula, it can be derived that one pixel corresponds to 0.215 mm.

2.2.5 Establishment of YOLOv7 model for seed germination detection

The seed quality detection methods such as germination and staining techniques are time-consuming and rely heavily on human intervention, which may lead to inaccurate results due to human error. In order to develop an automated and standardized method for detecting seed germination that is efficient, accurate, and reliable, the YOLOv7 (Wang et al., 2022b) object detection algorithm was selected in this study, which is one of the most widely used algorithms for object detection since its release in 2015 (Dewi et al., 2023). YOLOv7 is a real-time object detection algorithm (Soeb et al., 2023), which has evolved from YOLOv5 and has faster inference speed, improved detection accuracy, and reduced computational complexity. The algorithm consists of three main parts: the input layer, backbone layer, and output layer (Tang et al., 2023), and uses either a loss function with or without an auxiliary training head (Zhou et al., 2023).

The loss function is used to update the gradient loss during the training process (Cai et al., 2023). The YOLOv7 algorithm is evaluated using various metrics such as precision, mAP, recall, and F1 score (Zhao et al., 2023), and curves such as the F1-Confidence curve, precision-confidence curve, recall-confidence curve, and precision-recall curve are used to optimize the algorithm's performance and achieve the best balance between precision and recall.

This study utilized a self-built dataset of maize seeds, comprising images of seeds from various angles and sizes, each with corresponding labels in YOLO format. The data collection and preprocessing process was conducted using the same method as Mask R-CNN. The dataset used in this study consisted of a total of 7000 maize seeds. Among these, 4200 seeds were designated as training samples, 1400 seeds were allocated for the test sets, and the remaining seeds were assigned to the validation sets. To enhance the accuracy and robustness of the model, the YOLOv7.pt (https://github.com/WongKinYiu/yolov7) pretrained weights provided by the official website were employed for training. These weights were trained on a large-scale dataset, which can significantly reduce the training time while improving the training effect. The Adam optimizer, a widely used optimizer that can optimize at different learning rates, was used to update the model parameters during training. The parameters of the Adam optimizer were adjusted based on the size of the learning rate in the training process to achieve better training results. A batch size of 2 and a training iteration of 300 were used in this study.

3 Results and discussion

3.1 Seed germination result

The experimental results showed that the degree of seed aging was significantly correlated with the germination rate. As shown in

the Figure 4, on the seventh day of observation, all seeds that were not aged can germinate, and only a few seeds that aged for 2 days failed to do so. Most seeds that aged for 4 days still retained their viability, with only a few seeds that aged for 6 days able to germinate. Seeds that aged for 8 days experience almost complete mortality. Thus, it can be inferred that seed aging leads to a decline in the germination rate, and the more prolonged the aging process, the more apparent the decline in the germination rate.

3.2 Average spectrum

By analyzing the spectral curve features (Figures 5A, B), it is easy to observe that the spectral reflectance of both wavelength regions increased with the decreasing of maize seed viability, indicating that the light absorption capacity of maize tissue increases with the aging degree. The spectral curves are monotonic in the Vis-SWNIR region, with the average spectral curve gradually increasing in the 400-800 nm region and then slowly decreasing. However, in the LWNIR region, the spectral curve is more complex, capturing two distinct reflection peaks located around 1100 nm and 1300 nm, respectively. The former could potentially be associated with the presence of C-H bonds in lipids, while the latter could be described as a combination of the first overtone of N-H stretching along with the fundamental N-H in-plane bending and C-N stretching with N-H in-plane bending vibrations (Wang et al., 2022d).The spectral curve characteristics can be used to discriminate maize seeds with different germination potentials. As shown in Figure 6, the spectral data of maize seeds with different viability have similar trends in the Vis-SWNIR and LWNIR regions. However, in the Vis-SWNIR region, these curves are basically mixed together, making it difficult to distinguish clearly. In contrast, there are significant differences in the LWNIR region, which may be related to the breakdown of chemicals during the aging process of organisms. Nevertheless, some mixed situations still exist, indicating that it is difficult to distinguish the seeds with or without viability according to the average spectra of hyperspectral image.

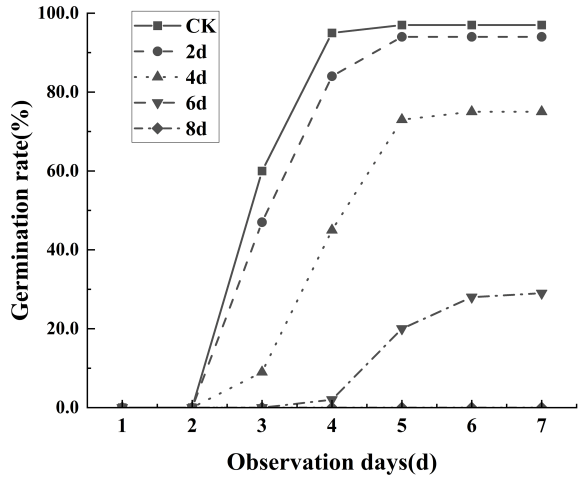


FIGURE 4
Germination levels of seeds at different aging times.

3.3 Key wavelength selection of maize seed viability

During the aging process of maize seeds, a series of changes occurs in the internal chemical substances (Xin et al., 2011), with the extent of these changes depending on the degree of seed vitality. These chemical substances include stored energy and nutrients, such as starch, proteins, and lipids (Xu et al., 2022). Proteins may undergo degradation, leading to the release of amino acids and structural damage to proteins. At the same time, the lipid content in the seed gradually oxidizes, resulting in lipid decomposition and the generation of free radicals, thereby affecting the seed's metabolism and viability. Additionally, starch gradually degrades into soluble sugars. This difference is the main reason for spectral changes during the aging process. After SG and SNV preprocessing, 18 and 11 characteristic bands were extracted from the Vis-SWNIR region and LWNIR region (Figures 5C, D). These characteristic bands were located at the peaks and valleys of the spectrum, reflecting the changes in water content and protein levels of the seeds.

3.4 Maize seed viability detection based on full-wavelength spectra and machine learning

By analyzing the classification accuracy obtained from SVM and Ensemble analysis, there was no significant difference between Vis-SWNIR and LWNIR regions in predicting seed viability (Table 3). However, KNN exhibited slightly higher accuracy with LWNIR, indicating its greater universality and better performance in detecting seed viability. However, due to the minimal differences between seeds with adjacent aging gradients (Feng et al., 2018), particularly those seeds that aged for 4 days and 6 days, these distinctions may not be immediately discernible, presenting a challenge in accurately determining the germination potential of seeds with similar levels of aging. The germination experiment also showed that the seeds with relatively mild aging did not have inherent germination trends and were easily misclassified by the prediction model. This discrepancy may arise from the fact that maize seeds may not exhibit overt phenotypic changes across different stages of aging (Wang et al., 2022e). However, in actuality, mRNA molecules associated with protein synthesis undergo oxidation through physiological mechanisms. More specifically, research unveiled significantly elevated expression levels of mature enzyme genes and ribosomal protein genes in embryonic roots and shoots as compared to other parts(Wang et al., 2022a). This obstruction hampers protein synthesis, consequently impeding the normal physiological functions of the seeds.

3.5 Maize seed viability detection based on key wavelength and 3DCNN model

After 70 training epochs on the Vis-SWNIR hyperspectral images, the accuracy of the training set has stabilized at a high level of 100% (Figure 6B), and the accuracy of the test set has also

10.3389/fpls.2023.1248598

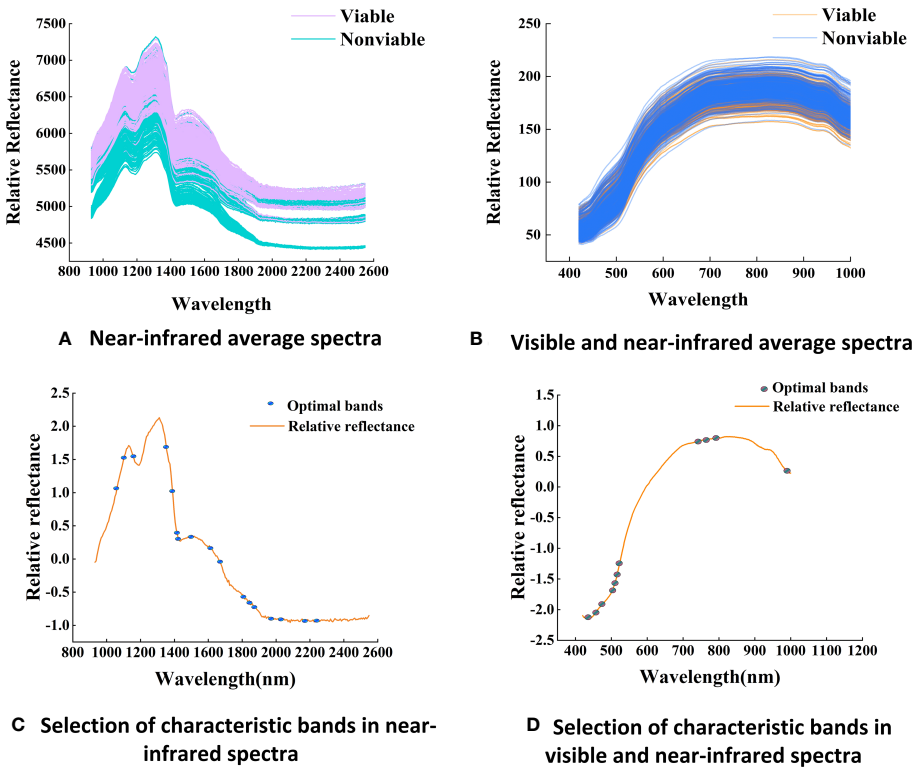


FIGURE 5
Average spectra and the distribution of optimal bands **(A)** Near-infrared average spectra, **(B)** Visible and near-infrared average spectra. **(C)** Selection of characteristic bands in near-infrared spectra, **(D)** Selection of characteristic bands in visible and near-infrared spectra.

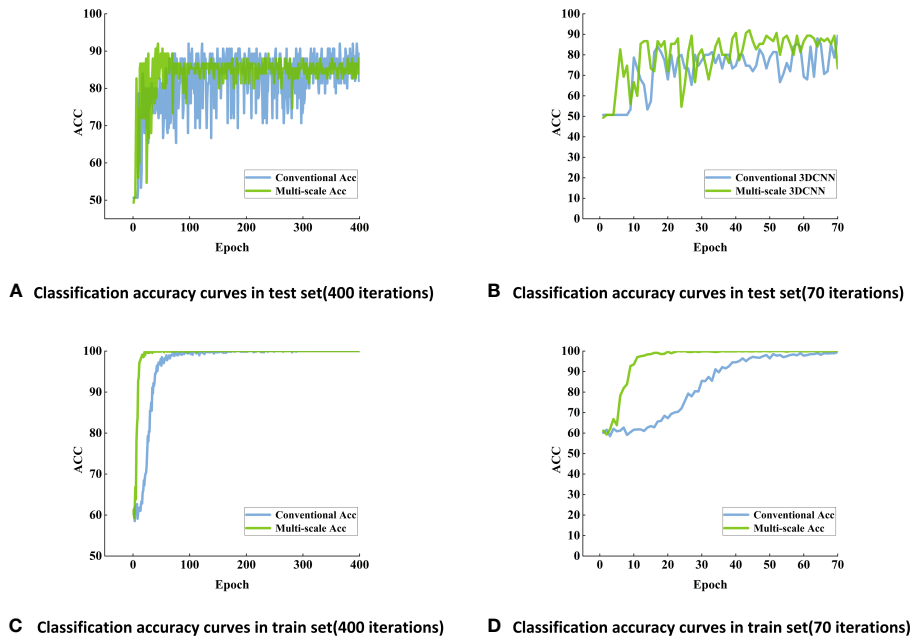


FIGURE 6
Classification accuracy curves of maize seed viability based on conventional 3DCNN models using Vis-SWNIR hyperspectral image **(A)** Classification accuracy curves in test set (400 iterations), **(B)** Classification accuracy curves in test set (70 iterations), **(C)** Classification accuracy curves in train set (400 iterations), **(D)** Classification accuracy curves in train set (70 iterations).

TABLE 3 The classification result of maize seed viability based on full-wavelength spectra and machine learning.

Models	Vis-SWNIR		LWNIR	
	Train set	Prediction set	Train set	Prediction set
SVM	89.3%	83.9%	84.0%	83.3%
KNN	72.0%	69%	85.3%	77.8%
Ensemble	92%	82.4%	85.3%	82.5%

reached its peak. In order to further validate the stability and robustness of the model, the number of training epochs was increased to 400. After 400 iterations, the accuracy of the training set remained at around 100%, while the accuracy of the test set remained at around 90% (Figures 6A, C).

By using 3DCNN to process the data, not only the spectral information was considered (Wu et al., 2021), but also the image information was integrated, making the evaluation of maize seed quality more comprehensive and accurate (Collins et al., 2021). Compared with machine learning methods that using all spectral bands as input data, the 3DCNN method only used few representative bands. Traditional machine learning methods tend to lose a lot of information, while the 3DCNN method used in this study can learn more complex features and achieved higher accuracy with fewer bands, with an average accuracy increase of 7 percentage points (Table 4). It was worth noting that 3DCNN performs better on the test set and converges faster, which indicated that 3DCNN was an effective method for seed viability detection and had advantages over machine learning classification method in dealing with such problems.

Conventional 3DCNN and multi-scale 3DCNN exhibit different characteristics. Traditional 3DCNN can achieve high accuracy, but they often exhibit slower convergence compared with multi-scale 3DCNN (Figure 6D). Multi-scale 3DCNN incorporated convolutional layers with different-sized kernels and pooling layers, allowing the network to process features of varying scales simultaneously (Lin et al., 2020). This enhanced the network's robustness and improved its tolerance to noise, distortion, and artifacts in the data, and ultimately led to a faster convergence. In addition, the stability of conventional 3DCNN may not be satisfactory and may exhibit some fluctuations and instability. In contrast, multi-scale 3DCNN perform better, possibly due to their utilization of multi-scale convolutional kernels, enabling them to extract more abundant feature information (Shi et al., 2021) (Figure 6A). Furthermore, the block-based method effectively increased the amount of data and helped to alleviate overfitting. In the final discrimination, this study adopted a majority principle labeling aggregation method to improve the discrimination accuracy (Table 4). To explore the optimal block size, several experiments were conducted, the input images were segmented into different block sizes, including 5 pixels ×5 pixels, 10 pixels ×10 pixels, and 20 pixels × 20 pixels. As shown in Table 4, the model achieved a relatively high overall accuracy when 5 pixels ×5 pixels was used. This suggested that the small blocks with 5 pixels ×5 pixels size can effectively capture more local features of the seedy and provides more discriminative information. Conversely, larger blocks may result in information blurring and confusion, thereby impacting the classification accuracy. Consequently, the block-based method with 5 pixels ×5 pixels was finally selected to enhances the detection accuracy of seed viability.

3.6 Maize seed germination detection based on YOLOv7 model

Figure 7 shows the detection results of germinated maize seeds using the YOLOv7 model, demonstrating its remarkable precision and recall rates of 99.7% and 99.0%, respectively. Additionally, the model achieves a mAP of 99% when applying an Intersection over Union (IoU) threshold of 0.5. Furthermore, the mAP, calculated across a range of IoU thresholds from 0.5 to 0.95, reaches a value of 71%.

In Table 5, the YOLOv7 model exhibits an impressive F1 score (The F1 score balances precision and recall, providing a comprehensive evaluation of model accuracy) of 0.99 on all target categories with a confidence threshold set at 0.663, highlighting its exceptional detection performance. Consequently, the YOLOv7 model can achieve both high precisions, accurately identifying

TABLE 4 The classification performance of the maize seed viability based on 3DCNN.

Block		Models			
		Multi-3DCNN		Conventional-3DCNN	
		Vis-SWNIR	SWNIR	Vis-SWNIR	SWNIR
Split	5 pixels×5 pixels	90.67%	90.67%	92.00%	88%
	10 pixels×10 pixels	92.00%	87.33%	92.00%	85.33%
	20 pixels×20 pixels	85.33%	79.00%	86.67%	78.67%
No-split		80.80%	78.50%	79.60%	77.50%

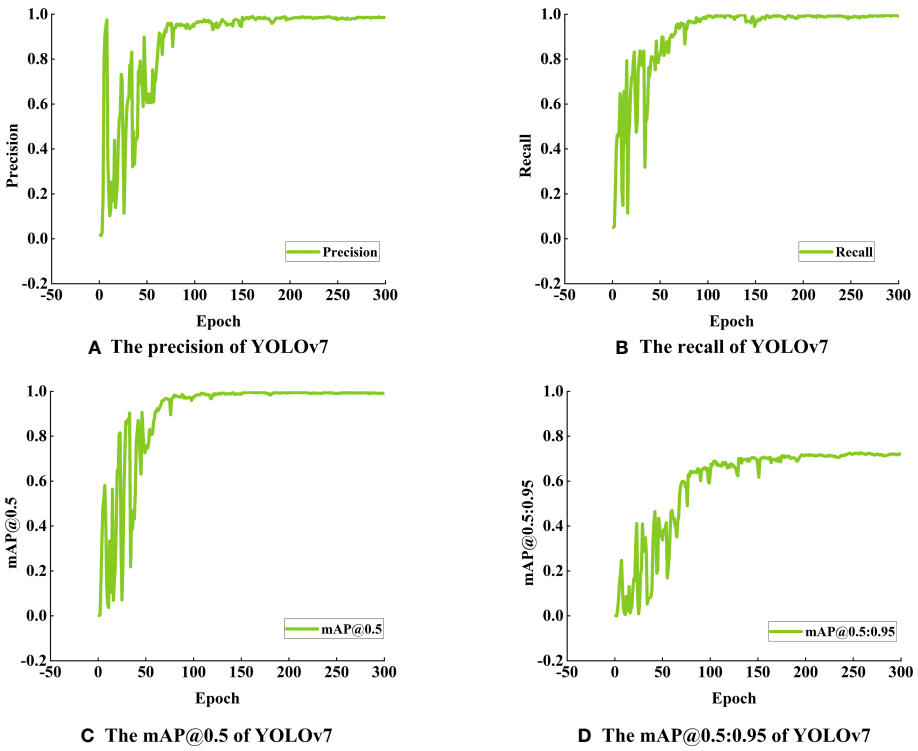


FIGURE 7
Detection performance of YOLOv7 model for maize seed germination **(A)** The precision of YOLOv7, **(B)** The recall of YOLOv7, **(C)** The mAP@0.5 of YOLOv7, **(D)** The mAP@0.5:0.95 of YOLOv7.

true positive predictions, and high recall, effectively capturing all relevant targets during detection. With a confidence threshold set to 0.896, the YOLOv7 model achieves a perfect precision accuracy of 100% for the target categories. This noteworthy precision metric showcases the model's ability to correctly identify true positive predictions among all the positive predictions made, indicating its reliability and precision in detecting target objects. The model impressively achieves a recall rate (The recall rate quantifies the model's ability to correctly identify positive targets) of 1.00 with a confidence threshold set to 0.000, indicating that it accurately detects all targets of all categories without any missed detections. This ideal performance underscores the model's high accuracy and proficiency in target detection tasks Additionally, the model exhibits an mAP (The mAP commonly used to evaluate object detection algorithms' accuracy and robustness) of 0.991 for all target categories when applying an Intersection over Union (IoU) threshold of 0.5. This further demonstrates the model's superior detection capabilities across various categories, affirming its exemplary performance.

$$R = \frac{TP}{TP + FN} \tag{7}$$

$$\mathrm{AP} = \int_0^1 P(R)dR \tag{8}$$

$$\mathrm{F1} = \frac{2 \star P \star R}{P + R} \tag{9}$$

$$\mathrm{P} = \frac{TP}{TP + FP} \tag{10}$$

$$\mathrm{mAP} = \frac{1}{n}\sum_{i=1}^{n} AP \tag{11}$$

TABLE 5 The detection result of YOLOv7 model for maize seed germination.

YOLOv7 Training Indicators					
All classes	**F1-confidence**	**F1**	0.99	**Confidence**	0.663
	Precision-confidence	**Precision**	1.00		0.896
	Recall-confidence	**Recall**	1.00		0.000
	Precision-recall	0.991 mAP@0.5			

In these formulas, True Positives (TP) represent the number of samples where the predicted label is positive and the actual label is also positive. T represents the total number of samples, and False Negatives (FN) indicate the number of samples where the predicted label is negative, but the actual label is positive. Similarly, False Positives (FP) represent the number of samples where the predicted label is positive, but the actual label is negative. Moreover, the area under the precision-recall (P-R) curve, denoted as AP, provides a measure of the model's performance.

Figure 8A is the confusion matrix of germinated maize seed based on YOLOv7 model, which provides a visual representation of the classification performance, showing the counts of true positive, true negative, false positive, and false negative predictions. The detection accuracy was 95% for germinated seeds and 99% for ungerminated seeds, respectively. Background FP refers to the number when the background is erroneously predicted as a target, fortunately there was no background area was incorrectly classified as a target in this study. Figure 8B shows the actual detection results of YOLOv7 for discriminating seed germination.

All indicators mean that the model can essentially replace manual observation for determining seed germination status. Therefore, although this method required some time and manpower for data annotation and training, the overall cost was much lower than manual operation, and can provide a reference for rapid detection of seed germination in crops. On the other hand, the algorithm suffered from the problem of duplicate detection in practical applications (Chen et al., 2023a), resulting in some seeds may be simultaneously labeled as germinated and non-germinated. This phenomenon may lead to a misclassification and reduce the practicality and reliability of the algorithm. Hence, future work will focus on improving the algorithm to solve the duplicate detection problem.

3.7 Maize seed bud length detection based on Mask R-CNN

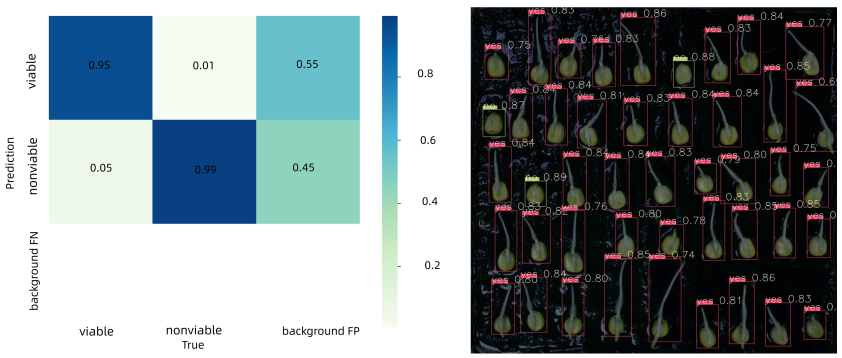
The Mask R-CNN model achieved an impressive mAP score of 0.9571, indicating its effectiveness and accuracy in detecting and localizing objects. The mAP is a widely used evaluation metric for object detection models, and a high mAP score indicates that the model performs well in both precision and recall, making it a reliable choice for seed germination analysis. Additionally, the loss value during training decreased significantly, stabilizing around 0.21 from an initial value of 2.61, which is a clear indication of the model's ability to learn and adapt effectively.

Figure 9A showcases a successful instance of skeleton extraction for maize seed germination, resulting in a clear main skeleton after removing branches, which allows for accurate calculation of the bud length. The detection results of bud length for germinated maize seeds, depicted in Figure 9B, demonstrate Mask R-CNN's impressive capability to accurately segment the seedlings, even when instances overlap or are occluded. This highlights the superiority of the Mask R-CNN model in instance segmentation tasks, making it a valuable tool for precise analysis of seed germination and growth.

Figures 9C, D shows the detection result of bud length with R-squared value of 0.98 and an RMSE of 1.64, demonstrating that the integration of Mask R-CNN model and skeleton extraction method could detect the bud length during seed germination accurately and rapidly. The R-squared value, also known as the coefficient of determination, is a statistical measure that indicates the proportion of the variance in the dependent variable (Bud length in this case) that can be explained by the independent variable (The predicted bud length). Meanwhile, RMSE quantifies the average magnitude of the differences between the predicted bud lengths and the actual observed bud lengths. It is worth mentioning that the bud length of germinated seeds is closely related to their viability (Adebisi et al., 2014). Therefore, the bud length of seeds can be obtained using this algorithm, and the relationship between bud length and viability can be further explored. This not only has important significance for agricultural production but also provides valuable insights for research in other biological fields.

$$SSR = \sum_{i=1}^{n}(Y_i - \hat{Y}_i)^2 \quad (12)$$

$$SST = \sum_{i=1}^{n}(Y_i - \bar{Y})^2 \quad (13)$$



A Confusion matrix **B** Image of detection results

FIGURE 8
Confusion matrix and detection results of germination maize seed based on YOLOv7 model **(A)** Confusion matrix, **(B)** Image of detection results.

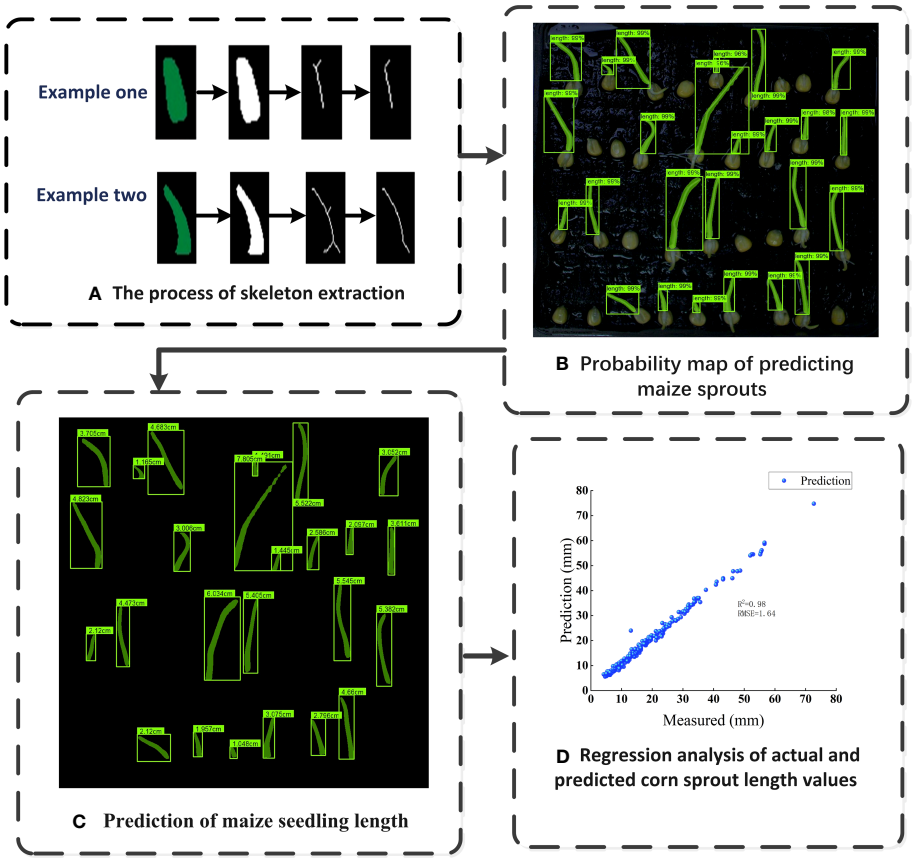


FIGURE 9
The bud length detection of germinated maize seeds **(A)** The process of skeleton extraction, **(B)** Probability map of predicting maize sprouts, **(C)** Prediction of maize seedling length, **(D)** Regression analysis of actual and predicted corn sprout length values.

$$RMSE = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(Y_i - \hat{Y}_i)^2} \quad (14)$$

$$R^2 = 1 - \frac{SSR}{SST} \quad (15)$$

In these formulas, SSR (Sum of Squares of Residuals) refers to the regression sum of squares, which represents the sum of squared differences between the predicted values and the true values. On the other hand, SST (Total Sum of Squares) stands for the total sum of squares, representing the sum of squared differences between the true values and their mean.Y_i refers to the actual value of the i-th observation, while $\hat{Y}_i$ represents the predicted value of the i-th observation from the regression model. And n denotes the sample size.

4 Conclusions

The rapid and successful detection of maize seed viability was achieved by leveraging HSI technology in combination with the multi-scale 3DCNN method. In seed viability detection, the 3DCNN method, which utilizes a limited number of representative spectral bands, was found to learn more complex features and achieve higher accuracy compared to using full-wavelength spectra and machine learning methods. By introducing the multi-scale 3DCNN model, the comprehensive consideration of both spectral and image information enabled a more comprehensive and accurate assessment of maize seed quality. Experimental results demonstrated that the adoption of small block sizes (5 pixels × 5 pixels) significantly improved the accuracy of seed viability detection. Furthermore, the YOLOv7 model and Mask R-CNN model were introduced for germination judgment and bud length detection of maize seeds. Both models exhibited outstanding performance in germination judgment and bud length detection, demonstrating excellent detection capabilities. Based on these exceptional detection results, a novel solution for the rapid detection of maize seed germination and bud length was provided. In brief, this study proposed a reliable and effective method for the evaluation of maize seed viability, providing valuable references for agricultural production and germplasm resource preservation.

Data availability statement

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

Author contributions

YF: Conceptualization, Data collection, Data analysis, Writing – original draft, Writing – review & editing. TA, GY, QW, WH, ZW: providing language help. CZ: Resources, Supervision. TX: Resources, Review-editing, Supervision. All authors contributed to the article and approved the submitted version.

Funding

This study was supported by China Postdoctoral Science Foundation (2022M720492), Postdoctoral Scientific Research Fund of Beijing Academy of Agricultural and Forestry Sciences (2022-ZZ-018), Beijing Postdoctoral Science Foundation (2023-117).

Conflict of interest

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

Publisher's note

References

Adebisi, M. A., Kehinde, T. O., Porbeni, J. B. O., Oduwaye, O. A., Biliaminu, K., and Akintunde, S. A. (2014). Seed and seedling vigour in tropical maize inbred lines. *Plant Breed. Seed Sci.* 67 (1), 87–102. doi: 10.2478/v10129-011-0072-4

Adegbuyi, E., and Burris, J. S. (1988). Field criteria used in determining the vigor of seed corn (Zea mays L.) as influenced by drying injury. *J. Agron. Crop Sci.* 161 (3), 171–177. doi: 10.1111/j.1439-037X.1988.tb00651.x

Aloupogianni, E., Ishikawa, M., Ichimura, T., Hamada, M., Murakami, T., Sasaki, A., et al. (2023). Effects of dimension reduction of hyperspectral images in skin gross pathology. *Skin Res. Technol.* 29 (2), e13270. doi: 10.1111/srt.13270

Ambrose, A., Kandpal, L. M., Kim, M. S., Lee, W.-H., and Cho, B.-K. (2016). High speed measurement of corn seed viability using hyperspectral imaging. *Infrared Phys. Technol.* 75, 173–179. doi: 10.1016/j.infrared.2015.12.008

An, T., Huang, W., Tian, X., Fan, S., Duan, D., Dong, C., et al. (2022). Hyperspectral imaging technology coupled with human sensory information to evaluate the fermentation degree of black tea. *Sensors Actuators B: Chem.* 366, 131994. doi: 10.1016/j.snb.2022.131994

Association, I.S.T (1999). *International rules for seed testing. Rules 1999.* (Zurich, Switzerland: International Seed Testing Association).

Bai, W., Zhao, X., Luo, B., Zhao, W., Huang, S., and Zhang, H. (2023). Research on wheat seed germination detection method utilizing YOLOv5. *Acta Agriculturae Zhejiangensis* 35 (02), 445–454. doi: 10.3969/j.issn.1004-1524.2023.02.22

Cai, L., Liang, J., Xu, X., Duan, J., and Yang, Z. (2023). Banana pseudostem visual detection method based on improved YOLOV7 detection algorithm. *Agronomy* 13 (4), 999. doi: 10.3390/agronomy13040999

Casado-García, Á., Domínguez, C., García-Domínguez, M., Heras, J., Inés, A., Mata, E., et al. (2019). CLoDSA: a tool for augmentation in classification, localization, detection, semantic segmentation and instance segmentation tasks. *BMC Bioinf.* 20 (1), 1–14. doi: 10.1186/s12859-019-2931-1

Chen, X., Li, F., and Chang, Q. (2023b). Combination of continuous wavelet transform and successive projection algorithm for the estimation of winter wheat plant nitrogen concentration. *Remote Sens.* 15 (4), 997. doi: 10.3390/rs15040997

Chen, J., Ma, B., Ji, C., Zhang, J., Feng, Q., Liu, X., et al. (2023a). Apple inflorescence recognition of phenology stage in complex background based on improved YOLOv7. *Comput. Electron. Agric.* 211, 108048. doi: 10.1016/j.compag.2023.108048

Cheng, T., Chen, G., Wang, Z., Hu, R., She, B., Pan, Z., et al. (2023). Hyperspectral and imagery integrated analysis for vegetable seed vigor detection. *Infrared Phys. Technol.* 131, 104605. doi: 10.1016/j.infrared.2023.104605

Collins, T., Maktabi, M., Barberio, M., Bencteux, V., Jansen-Winkeln, B., Chalopin, C., et al. (2021). Automatic recognition of colon and esophagogastric cancer with machine learning and hyperspectral imaging. *Diagnostics* 11 (10), 1810. doi: 10.3390/diagnostics11101810

Cong, P., Li, S., Zhou, J., Lv, K., and Feng, H. (2023). Research on instance segmentation algorithm of greenhouse sweet pepper detection based on improved mask RCNN. *Agronomy* 13 (1), 196. doi: 10.3390/agronomy13010196

Cortes, C., and Vapnik, V. (1995). Support-vector networks. Support-vector networks. *Machine Learning* 20 (3), 273–297. doi: 10.1007/BF00994018

Cui, H., Bing, Y., Zhang, X., Wang, Z., Li, L., and Miao, A. (2022). Prediction of maize seed vigor based on first-order difference characteristics of hyperspectral data. *Agronomy* 12 (8), 1899. doi: 10.3390/agronomy12081899

Cui, H., Cheng, Z., Li, P., and Miao, A. (2020). Prediction of sweet corn seed germination based on hyperspectral image technology and multivariate data regression. *Sensors* 20 (17), 4744. doi: 10.3390/s20174744

de Almeida, V. E., de Araújo Gomes, A., de Sousa Fernandes, D. D., Goicoechea, H. C., Galvão, R. K. H., and Araújo, M. C. U. (2018). Vis-NIR spectrometric determination of Brix and sucrose in sugar production samples using kernel partial least squares with interval selection based on the successive projections algorithm. *Talanta* 181, 38–43. doi: 10.1016/j.talanta.2017.12.064

Dewi, C., Chen, A. P. S., and Christanto, H. J. (2023). Deep learning for highly accurate hand recognition based on yolov7 model. *Big Data Cogn. Computing* 7 (1), 53. doi: 10.3390/bdcc7010053

Fan, S., Li, C., Huang, W., and Chen, L. (2018). Data fusion of two hyperspectral imaging systems with complementary spectral sensing ranges for blueberry bruising detection. *Sensors* 18 (12), 4463. doi: 10.3390/s18124463

Feng, L., Zhu, S., Zhang, C., Bao, Y., Feng, X., and He, Y. (2018). Identification of maize kernel vigor under different accelerated aging times using hyperspectral imaging. *Molecules* 23 (12), 3078. doi: 10.3390/molecules23123078

Fu, Y., Xia, Y., Zhang, H., Fu, M., Wang, Y., Fu, W., et al. (2023). Skeleton extraction and pruning point identification of jujube tree for dormant pruning using space colonization algorithm. *Front. Plant Sci.* 13. doi: 10.3389/fpls.2022.1103794

Gaikwad, J., Triki, A., and Bouaziz, B. (2019). Measuring morphological functional leaf traits from digitized herbarium specimens using traitEx software. *Biodiversity Inf. Sci. Standards* 3, e37091. doi: 10.3897/biss.3.37091

Ge, Z., Cao, G., Li, X., and Fu, P. (2020). Hyperspectral image classification method based on 2D–3D CNN and multibranch feature fusion. *IEEE J. Selected Topics Appl. Earth Observations Remote Sens.* 13, 5776–5788. doi: 10.1109/jstars.2020.3024841

Ghaderizadeh, S., Abbasi-Moghadam, D., Sharifi, A., Zhao, N., and Tariq, A. (2021). Hyperspectral image classification using a hybrid 3D-2D convolutional neural networks. *IEEE J. Selected Topics Appl. Earth Observations Remote Sens.* 14, 7570–7588. doi: 10.1109/jstars.2021.3099118

Han, X., Jiang, Z., Liu, Y., Zhao, J., Sun, Q., and Li, Y. (2022). A spatial–spectral combination method for hyperspectral band selection. *Remote Sens.* 14 (13), 3217. doi: 10.3390/rs14133217

He, K., Gkioxari, G., Dollár, P., and Girshick, R. (2017). Mask r-cnn in *Proceedings of the IEEE international conference on computer vision*. 2961–2969.

Köppen, M. (2000). The curse of dimensionality, in *5th online world conference on soft computing in industrial applications*. (WSC5). 1, 4–8.

Li, Y., Zhang, H., and Shen, Q. (2017). Spectral–spatial classification of hyperspectral imagery with 3D convolutional neural network. *Remote Sens.* 9 (1), 67. doi: 10.3390/rs9010067

Liang, X.-y., Guo, F., Feng, Y., Zhang, J.-L., Yang, S., Meng, J.-J., et al (2020). Single-seed sowing increased pod yield at a reduced seeding rate by improving root physiological state of Arachis hypogaea. *J. Integr. Agric.* 19 (4), 1019–1032. doi: 10.1016/S2095-3119(19)62712-7

Lin, B., Zhang, S., and Bao, F. (2020). "Gait recognition with multiple-temporal-scale 3D convolutional neural network," in *Proceedings of the 28th ACM International Conference on Multimedia*. (Seattle, WA, USA: Association for Computing Machinery).

Liu, C., Chu, Z., Weng, S., Zhu, G., Han, K., Zhang, Z., et al. (2022). Fusion of electronic nose and hyperspectral imaging for mutton freshness detection using input-

modified convolution neural network. *Food Chem.* 385, 132651. doi: 10.1016/j.foodchem.2022.132651

Long, Y., Wang, Q., Tang, X., Tian, X., Huang, W., and Zhang, B. (2022). Label-free detection of maize kernels aging based on Raman hyperspcectral imaging techinique. *Comput. Electron. Agric.* 200, 107229. doi: 10.1016/j.compag.2022.107229

Masood, M., Nazir, T., Nawaz, M., Javed, A., Iqbal, M., and Mehmood, A. (2021). Brain tumor localization and segmentation using mask RCNN. *Front. Comput. Sci.* 15 (6). doi: 10.1007/s11704-020-0105-y

Pang, L., Wang, J., Men, S., Yan, L., and Xiao, J. (2021). Hyperspectral imaging coupled with multivariate methods for seed vitality estimation and forecast for Quercus variabilis. *Spectrochimica Acta Part A: Mol. Biomolecular Spectrosc.* 245, 118888. doi: 10.1016/j.saa.2020.118888

Shen, R., Zhen, T., and Li, Z. (2023). Segmentation of unsound wheat kernels based on improved mask RCNN. *Sensors* 23 (7), 3379. doi: 10.3390/s23073379

Shi, W., Du, C., Gao, B., and Yan, J. (2021). Remote sensing image fusion using multi-scale convolutional neural network. *J. Indian Soc. Remote Sens.* 49 (7), 1677–1687. doi: 10.1007/s12524-021-01353-2

Soeb, M. J. A., Jubayer, M. F., Tarin, T. A., Al Mamun, M. R., Ruhad, F. M., Parven, A., et al (2023). Tea leaf disease detection and identification based on YOLOv7 (YOLO-T). *Sci. Rep.* 13 (1), 6078. doi: 10.1038/s41598-023-33270-4

Suksungworn, R., Sanevas, N., Wongkantrakorn, N., Fangern, N., Vajrodaya, S., and Duangsrisai, S. (2021). Phytotoxic effect of Haldina cordifolia on germination, seedling growth and root cell viability of weeds and crop plants. *NJAS: Wageningen J. Life Sci.* 78 (1), 175–181. doi: 10.1016/j.njas.2016.05.008

Sun, W., and Du, Q. (2019). Hyperspectral band selection: A review. *IEEE Geosci. Remote Sens. Magazine* 7 (2), 118–139. doi: 10.1109/mgrs.2019.2911100

Sun, K., Wang, A., Sun, X., and Zhang, T. (2022). Hyperspectral image classification method based on M-3DCNN-Attention. *J. Appl. Remote Sens.* 16 (02), 026507. doi: 10.1117/1.Jrs.16.026507

Tang, F., Yang, F., and Tian, X. (2023). Long-distance person detection based on YOLOv7. *Electronics* 12 (6), 1502. doi: 10.3390/electronics12061502

Tian, X., Zhang, C., Li, J., Fan, S., Yang, Y., and Huang, W. (2021). Detection of early decay on citrus using LW-NIR hyperspectral reflectance imaging coupled with two-band ratio and improved watershed segmentation algorithm. *Food Chemistry* 360, 130077. doi: 10.1016/j.foodchem.2021.130077

Triki, A., Bouaziz, B., Gaikwad, J., and Mahdi, W. (2021). Deep leaf: Mask R-CNN based leaf detection and segmentation from digitized herbarium specimen images. *Pattern Recognition Lett.* 150, 76–83. doi: 10.1016/j.patrec.2021.07.003

Wang, C. -Y., Bochkovskiy, A., and Liao, H. -Y. M. (2022b). "YOLOv7: Trainable bag-of-freebies sets new state-of-the-art for real-time object detectors", in: *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition)*, 7464–7475. doi: 10.48550/arXiv.2207.02696

Wang, Z., Huang, W., Tian, X., Long, Y., Li, L., and Fan, S. (2022e). Rapid and non-destructive classification of new and aged maize seeds using hyperspectral image and chemometric methods. *Front. Plant Sci.* 13. doi: 10.3389/fpls.2022.849495

Wang, W., Huang, W., Yu, H., and Tian, X. (2022d). Identification of maize with different moldy levels based on catalase activity and data fusion of hyperspectral images. *Foods* 11 (12), 1727. doi: 10.3390/foods11121727

Wang, S., Sun, G., Zheng, B., and Du, Y. (2021). A crop image segmentation and extraction algorithm based on mask RCNN. *Entropy* 23 (9), 1160. doi: 10.3390/e23091160

Wang, J., Yan, L., Wang, F., Qi, S., and Jin, X.-B. (2022c). SVM classification method of waxy corn seeds with different vitality levels based on hyperspectral imaging. *J. Sensors* 1–13. doi: 10.1155/2022/4379317

Wang, B., Yang, R., Ji, Z., Zhang, H., Zheng, W., Zhang, H., et al. (2022a). Evaluation of biochemical and physiological changes in sweet corn seeds under natural aging and artificial accelerated aging. *Agronomy* 12 (5), 1028. doi: 10.3390/agronomy12051028

Wu, H., Li, D., Wang, Y., Li, X., Kong, F., and Wang, Q. (2021). Hyperspectral image classification based on two-branch spectral–spatial-feature attention network. *Remote Sens.* 13 (21), 4262. doi: 10.3390/rs13214262

Xin, X., Lin, X.-H., Zhou, Y.-C., Chen, X.-L., Liu, X., and Lu, X.-X. (2011). Proteome analysis of maize seeds: the effect of artificial ageing. *Physiologia Plantarum* 143 (2), 126–138. doi: 10.1111/j.1399-3054.2011.01497.x

XingJia, T., PengChang, Z., ZongBen, X., BingLiang, H., and Lakshmanna, K. (2022). Calligraphy and painting identification 3D-CNN model based on hyperspectral image MNF dimensionality reduction. *Comput. Intell. Neurosci.* 2022, 1–19. doi: 10.1155/2022/1418814

Xu, Y., Ma, P., Niu, Z., Li, B., Lv, Y., Wei, S., et al. (2022). Effects of artificial aging on physiological quality and cell ultrastructure of maize (Zea mays L.). *Cereal Res. Commun.* 51, 615–626. doi: 10.1007/s42976-022-00328-4

Yasmin, J., Ahmed, M. R., Wakholi, C., Lohumi, S., Mukasa, P., Kim, G., et al. (2022). Near-infrared hyperspectral imaging for online measurement of the viability detection of naturally aged watermelon seeds. *Front. Plant Sci.* 13. doi: 10.3389/fpls.2022.986754

Zhang, S. (2022). Challenges in KNN classification. *IEEE Trans. Knowledge Data Eng.* 34 (10), 4663–4675. doi: 10.1109/tkde.2021.3049250

Zhang, T., Fan, S., Xiang, Y., Zhang, S., Wang, J., and Sun, Q. (2020). Non-destructive analysis of germination percentage, germination energy and simple vigour index on wheat seeds during storage by Vis/NIR and SWIR hyperspectral imaging. *Spectrochimica Acta Part A: Mol. Biomolecular Spectrosc.* 239, 118488. doi: 10.1016/j.saa.2020.118488

Zhao, J., Ma, Y., Yong, K., Zhu, M., Wang, Y., Luo, Z., et al. (2022). Deep-learning-based automatic evaluation of rice seed germination rate. *J. Sci. Food Agric.* 103 (4), 1912–1924. doi: 10.1002/jsfa.12318

Zhao, J., Ma, Y., Yong, K., Zhu, M., Wang, Y., Wang, X., et al. (2023). Rice seed size measurement using a rotational perception deep learning model. *Comput. Electron. Agric.* 205, 107583. doi: 10.1016/j.compag.2022.107583

Zhao, W., Chellappa, R., and Phillips, P. J. (1999). Subspace linear discriminant analysis for face recognition. University of Maryland at College Park, USA: Citeseer.

Zhou, J., Zhang, Y., and Wang, J. (2023). RDE-YOLOv7: an improved model based on YOLOv7 for better performance in detecting dragon fruits. *Agronomy* 13 (4), 1042. doi: 10.3390/agronomy13041042

frontiers | Frontiers in Plant Science

TYPE Original Research
PUBLISHED 08 September 2023
DOI 10.3389/fpls.2023.1254548

OPEN ACCESS

EDITED BY
Xi Tian,
Beijing Academy of Agriculture and Forestry Sciences, China

REVIEWED BY
Yanru Zhao,
Henan Polytechnic University, China
Longzhe Quan,
Anhui Agricultural University, China

*CORRESPONDENCE
Fu Zhang
zhangfu30@126.com
Sanling Fu
fusanling@126.com

RECEIVED 07 July 2023
ACCEPTED 16 August 2023
PUBLISHED 08 September 2023

CITATION
Zhang F, Zhang F, Wang S, Li L, Lv Q, Fu S, Wang X, Lv Q and Zhang Y (2023) Hyperspectral imaging combined with CNN for maize variety identification. *Front. Plant Sci.* 14:1254548. doi: 10.3389/fpls.2023.1254548

Hyperspectral imaging combined with CNN for maize variety identification

Fu Zhang[1,2]*, Fangyuan Zhang[1], Shunqing Wang[1], Lantao Li[3], Qiang Lv[4], Sanling Fu[5]*, Xinyue Wang[1], Qingfeng Lv[1] and Yakun Zhang[1]

[1]College of Agricultural Equipment Engineering, Henan University of Science and Technology, Luoyang, China, [2]Collaborative Innovation Center of Machinery Equipment Advanced Manufacturing of Henan Province, Henan University of Science and Technology, Luoyang, China, [3]College of Resources and Environment, Henan Agricultural University, Zhengzhou, China, [4]College of Agriculture / Peon, Henan University of Science and Technology, Luoyang, China, [5]College of Physical Engineering, Henan University of Science and Technology, Luoyang, China

Introduction: As the third largest food crop in the world, maize has wide varieties with similar appearances, which makes identification difficult. To solve the problem of identification of hybrid maize varieties, a method based on hyperspectral image technology combined with a convolutional neural network (CNN) is proposed to identify maize varieties.

Methods: In this study, 735 maize seeds from seven half-parent hybrid maize varieties were regarded as the research object. The maize seed images in the range of 900 ~ 1700nm were obtained by hyperspectral image acquisition system. The region of interest (ROI) of the embryo surface was selected, and the spectral reflectance of maize seeds was extracted. After Savitzky-Golay (SG) Smoothing pretreatment, Maximum Normalization (MN) pretreatment was performed. The 56 feature wavelengths were selected by Competitive Adaptive Reweighting Algorithm (CARS) and Successive Projection Algorithm (SPA). And the 56 wavelengths were mapped to high-dimensional space by high-dimensional feature mapping and then reconstructed into three-dimensional image features. A five-layer convolution neural network was used to identify three-dimensional image features, and nine (SG+MN)-(CARS+SPA)-CNN maize variety identification models were established by changing the input feature dimension and the depth factor size of the model layer.

Results and Discussion: The results show that the maize variety classification model works best, when the input feature dimension is 768 and the layer depth factor d is 1.0. At this point, the model accuracy of the test set is 96.65% and the detection frame rate is1000 Fps/s in GPU environment, which can realize the rapid and effective non-destructive detection of maize varieties. This study provides a new idea for the rapid and accurate identification of maize seeds and seeds of other crops.

KEYWORDS

hyperspectral imaging technology, maize, high dimensional feature mapping, convolution neural network, non-destructive testing

1 Introduction

As one of the three major food crops in the world, maize has a wide cultivation area, large yield and strong adaptability, which is of great strategic significance to the economic development and social stability of China (Feng et al., 2022). In the actual agricultural production process, cultivating suitable maize varieties is the crucial aspect to achieving high yields. Different maize seeds are easily confused due to various sorts and similar appearance, which brings great inconvenience to farmers in purchasing varieties and market supervision (Tu et al., 2021). Therefore, it is of great significance and application value to achieve rapid, accurate and efficient identification of maize varieties.

Traditional seed identification methods mainly include manual detection methods, chemical identification methods and so on, which have some defects such as intense subjectivity, great destructiveness and complex operation process. And they are challenging to meet the requirements of modern agriculture for non-destructive and rapid seed production (Wang et al., 2021; Wang and Wang 2021; Huang et al., 2022). Hyperspectral imaging technology combines the advantages of image and spectral technology, which can simultaneously reflect the image information and spectral information of external characteristics, internal physical structure and chemical composition of samples to be tested. So hyperspectral imaging technology is widely used in non-destructive testing research on crop seed varieties, quality and vitality (Wu et al., 2021; Wu et al., 2022; Yang et al., 2022). Huang et al. (2016b) used hyperspectral imaging technology to establish a PLS-SVM model to identify four different years of maize seeds, and the identification accuracy rate reached 94.4%. Fu et al. (2022) identified four maize varieties based on hyperspectral imaging technology, and the accuracy of the SSAE-CS-SVM model test set reached 95.81%. Wang et al. (2022) used hyperspectral imaging technology to establish a fusion model based on dual-band ratio image and texture features to realize efficient non-destructive identification of maize seeds of four different years, and the accuracy rate of prediction set was 97.5%. Chivasa et al. (2019) developed a PLS-DA model based on multi-temporal hyperspectral data and multivariate techniques to identify 25 maize varieties at specific phenological stages. Tu et al. (2022) used hyperspectral imaging technology combined with machine learning to realize non-destructive identification of 10 related maize varieties. Huang et al. (2016a) established the LS-SVM maize variety classification model based on hyperspectral imaging combined with spectral features and fusion with image features to identify 17 maize varieties with a test accuracy of more than 90%. Wu et al. (2016) collected hyperspectral image data of four maize varieties based on NIR hyperspectral technology, and established the SPA-PLS-DA classification model to realize non-destructive identification of maize varieties. The accuracy of the modeling set and prediction set reached 78.5% and 70.8%, respectively. Shao et al. (2019) collected hyperspectral images of three varieties of maize based on hyperspectral imaging system, screened characteristic bands by Boruta algorithm, and established a random forest classification model, with an accuracy rate of 78.30%. Sun M et al. (2022) modeled and analyzed wheat seeds of different seven years based on hyperspectral imaging technology, and predicted wheat seeds of other four years with an accuracy rate of 100%. Zhang et al. (2019) used hyperspectral image technology to obtain hyperspectral image spectral information of the wheat varieties mainly planted in Henan Province, and established PCA-SVM classification model, which identified three wheat varieties with an accuracy rate of over 95%. Sun Y et al. (2022) modeled and analyzed spectral information of moldy and non-moldy grains of different wheat varieties collected by hyperspectral imaging technology, and the prediction accuracy of SPA-SVM model for moldy grains was more than 98%.

Existing research shows that a large number of scholars at home and abroad have carried out research on crop seed variety identification, most of which are based on two methods: hyperspectral image information combined with deep learning and modeling based on spectral data. The method based on hyperspectral image mainly applies image features to identify seed varieties, which is suitable for identifying seed varieties with obvious shape and texture differences. While most seeds have no noticeable appearance differences, therefore, this method is difficult to be widely used in the identification of crop seed varieties in practice. The modeling methods based on spectral information are divided into two steps: feature band extraction and model building. The feature bands are mostly extracted by single extraction method such as SPA or CARS, which has some problems such as incomplete feature band extraction and lack of effective information. In addition, traditional machine learning models such as SVM, PLS and PCA are primarily used in modeling methods, which have the disadvantages of low accuracy and poor robustness. The convolutional neural network, as a kind of forward feedback network, can automatically learn the features in the image with higher accuracy and efficiency. Hybrid maize varieties are similar in appearance and not easily distinguishable, and subtle differences in the content of internal substances cause significant differences in yield, insect resistance, disease resistance, stress resistance and other indicators. Based on this situation, this paper was conducted with seven hybrid maize varieties as the research object, using SPA and CARS mixed feature band extraction method to improve the utilization rate of effective feature information, and building a convolution neural network (CNN) model based on data reshaping to achieve accurate identification of maize varieties.

2 Materials and methods

2.1 Test materials

The maize seed samples used in the experiment are all from the experimental field of Yanshi District, Luoyang City, Henan Province. Seven half-parent hybrid maize varieties with good appearance, uniform color and no mechanical damage were manually selected, marked as categories 0, 1, 2, 3, 4, 5 and 6, respectively. As shown in Figure 1, there were 105 seeds in each category, with a total of 735 test samples.

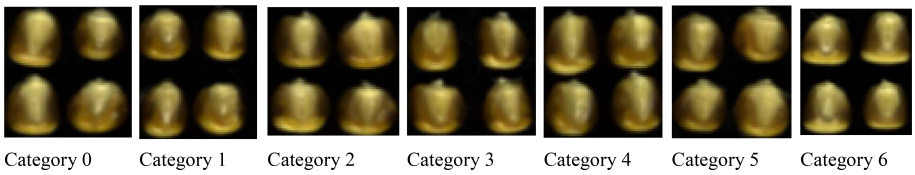


FIGURE 1
Maize seed sample.

2.2 Instruments and equipment

The hyperspectral imaging system used in this test consists of a hyperspectral camera (SPECIM FX17, Finland), a computer (Dell), a mobile platform, a sample tray (40cm × 20cm) and six halogen lamps. To eliminate the effect of ambient light, maize seed spectra were collected in a dark box, as shown in Figure 2.

The parameters of the hyperspectral imaging system were set as follows: wavelength range is 900 ~ 1700nm, spectral resolution is 8nm, the number of bands is 224, spatial sampling resolution is 640px/line, exposure time is 8.5 ms, the frame rate is 50Hz and platform moving speed is 22.43 mm/s. Hyperspectral data of maize seeds were obtained by using Lumo Scanner software. The data analysis software is as follows: ENVI 5.3, The Unscrambler X10.4, Excel 2019, Origin 2018, MATLAB R2018b and so on.

2.3 Hyperspectral image acquisition and correction

Maize seed embryos are rich in nutrients such as starch and protein, so the embryo surface image information of the sample was collected in this experiment (Wang et al., 2019). To ensure the accuracy of collected data and prevent maize seeds from moving on the mobile platform, the samples were placed on sticky black paper with their embryo face up. As shown in Figure 3, the images of 105 maize seeds of one variety were collected at a time, and a total of 735 single maize seed samples images were collected in the experiment.

Hyperspectral image is easily affected by nonlinear factors such as uneven distribution of light sources and dark current. To enhance the stability and reliability of the image, dark and white reference calibration images were used to correct the original hyperspectral image. The hyperspectral system was preheated for 30 minutes, the whiteboard (reflectivity 99%) was scanned and an all-white calibration image was recorded as I_w, the lens cover was installed to collect all-black image which was recorded as I_d, and finally the original image of maize sample was photographed and recorded as I_{raw}. And the corrected image I is obtained by black-and-white correction with ENVI 5.3 software.

$$I = \frac{I_{raw} - I_d}{I_w - I_d} \quad (1)$$

After image correction, to reduce the influence of uneven distribution of chemical components in seeds, the largest possible rectangular ROI region was selected in the center of each seed sample by ENVI 5.3 software, and the average of the spectra of all pixel points within the ROI region was taken as the average spectrum of the sample (Feng et al., 2012). The original spectral average reflectance curve in the wavelength range of 935.61 ~ 1720.23 nm is shown in Figure 4. Due to both ends of the collected spectrum with low signal-to-noise ratio, the areas with considerable noise of spectral signal are eliminated, and the spectral data range of 949.43 ~ 1709.49 nm wavelength are selected for analysis and modeling.

2.4 Spectral preprocessing and feature wavelength selection

The noise, background and other useless interference information mixed in the acquisition of spectra affected the accuracy and stability of spectral data analysis and modeling, so it is necessary to preprocess the data before modeling to reduce the interference of irrelevant information and improve the modeling

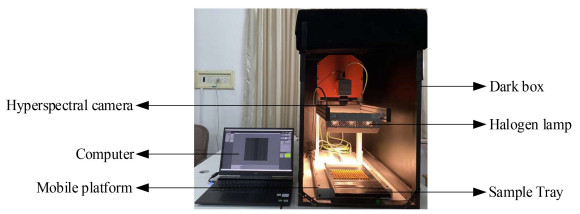


FIGURE 2
Hyperspectral image acquisition system.

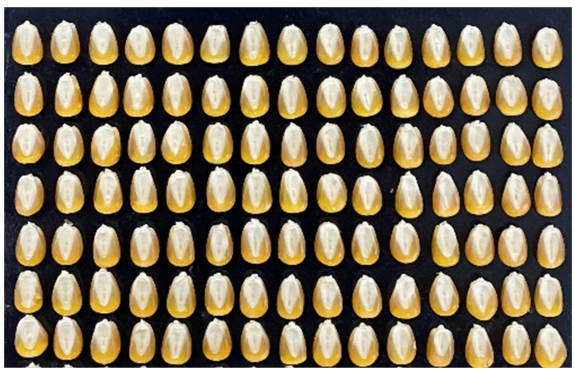


FIGURE 3
Schematic diagram of maize grain placement.

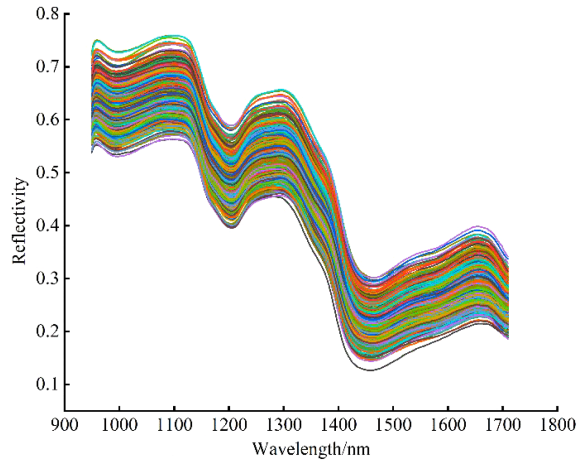


FIGURE 4
Reflectance curves of original spectrum.

accuracy. In this study, Savitzky-Golay Smoothing (SG) and Maximum Normalization (MN) were used to preprocess the data. Firstly, the number of smoothing points was set to 3, and the spectral data was pretreated by SG to improve the smoothness of the spectral curve. After that, the spectral data were mapped to the [0, 1] interval by MN, and the data unit restriction was removed to eliminate the errors caused by different magnitudes. The pretreated spectral average reflectivity curves are shown in Figure 5.

The Successive projections algorithm (SPA) was used to extract the feature bands from the pretreated spectrum, the maximum number of wavelengths was set to 20, and five wavelength variables were extracted, as shown in Figure 6. As can be seen from Figure 6A, with the increase of the number of variables, the root mean square error (RMSE) value shows a trend of sharp drop at first and then slow down. When the number of variables is 5, the RMSE no longer decreases significantly and the RMSE value is 1.7221 at this time. After that, although the REMSE value decreases, too many dependent variables will increase the computation and

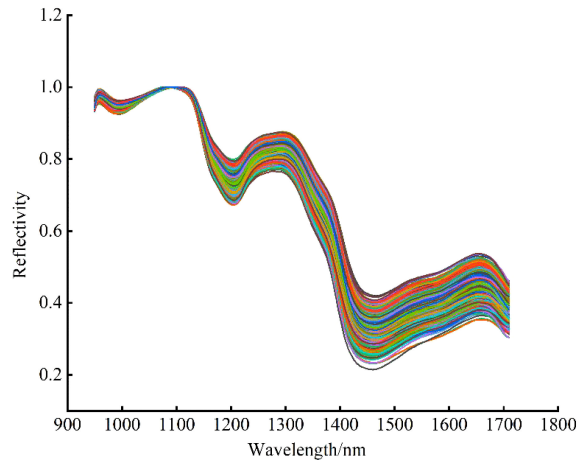


FIGURE 5
Spectral average reflectance curves after pretreatment.

complexity of the model, so five variables are selected as the final characteristic wavelengths.

When using Competitive adaptive reweighted sampling (CARS) to extract feature wavelengths, the 5-fold cross-validation method was selected, and the number of Monte Carlo samples was set to 50, as shown in Figure 7. Figure 7A shows that the number of CARS extracted feature wavelengths decreases sharply at first and then decreases slowly with the increase of sampling times, which shows the process from coarse to fine selection of feature wavelengths extracted by CARS. Figure 7B shows that the root mean square error of cross-validation (RMSECV) decreases slowly at the beginning of the iteration because the useless information bands are eliminated. And the RMSECV increases gradually after the 24th sampling, which indicates that the over-selection of feature wavelengths by CARS occurs after the 24th sampling and sensitive wavelength variables containing valid information are eliminated, resulting in the decrease of model prediction accuracy and the increase of RMSECV value; Figure 7C indicates that the RMSECV value is the smallest at the 6th and 16th sampling, when 52 characteristic wavelengths are extracted.

To solve the problem of missing effective information in the single extraction of feature variables by SPA and CARS, the feature wavelengths extracted by the two methods were taken as a concatenated set in this study, and a total of 56 feature wavelengths were preferentially selected.

2.5 Division of training set and test set

In this experiment, 735 samples were divided into training sets and test sets according to the ratio of 2: 1, where each category of training sets and test sets were 70 and 35 respectively. And seven categories of training sets and test sets were 490 and 245 respectively to analyze and calculate the discrimination accuracy of model training sets and test sets.

3 Model construction

3.1 Establishment of maize variety identification model

To solve the problem that the Convolutional Neural Networks (CNN) cannot directly process the feature band data, the maize seed feature wavelength data was mapped to the high-dimensional features. Then the mapped feature wavelength data was reshaped into high-dimensional image features making the CNN processable for the reshaped data. The overall network structure is shown in Figure 8. The CNN model consists of three parts: data reconstruction, convolution layer extraction and result prediction. In the data reconstruction part, the feature bands of maize seeds are mapped into high-dimensional features with different sizes, and then reshaped into image shapes. Considering the dimension of maize seed characteristic band, it is not easy to build the convolution layers too deep to avoid overfitting and poor robustness of the model. Therefore, a 5-layer CNN maize variety

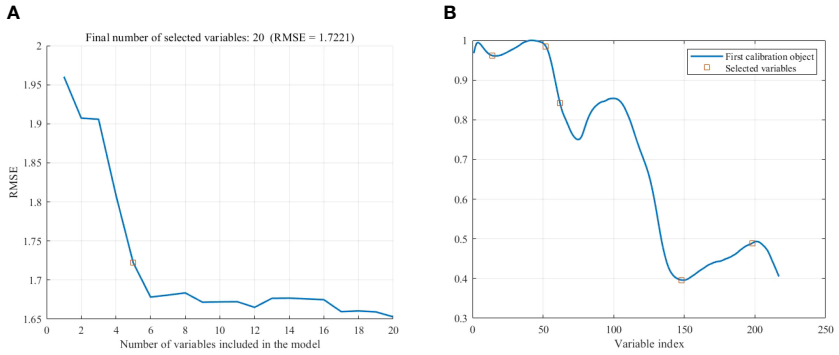


FIGURE 6
Extraction of feature wavelength by SPA. **(A)** Number of variables. **(B)** Location of variables.

discrimination model was constructed to improve the model generalization performance and reduce the redundancy effect of the model on spatial features. The specific model parameters are shown in Table 1. It can be seen from Table 1 that the overall maize variety identification model is built by 3×3 standard convolution. To improve the spatial feature extraction effect of the model on maize seed feature bands, the sampling method of raising dimension first and then reducing dimension is adopted to fuse the features effectively. In order to explore the influence of model layers depth on the sampling effect of maize seed characteristic band, three common scaling factors (d), 1.25, 1.0 and 0.75 were used to scale the layers of maize variety identification model to different degrees. And the related parameters are listed in Table 1. In addition, to explore the influence of different high-dimensional feature resolutions on the adaptability of maize variety identification model and find the best adaptability resolution of the model, three different feature mapping relationships of 192, 768 and 3072 were used to generate three corresponding spatial feature

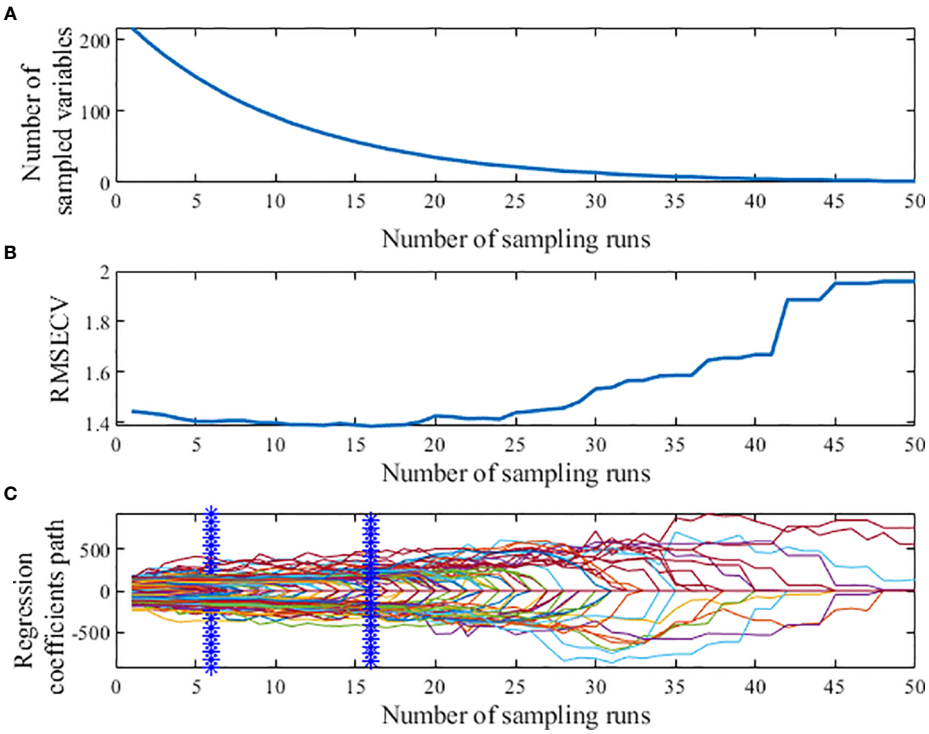


FIGURE 7
Process of extracting characteristic wavelength by CARS. **(A)** Number of preferred characteristic wavelength variables. **(B)** The root mean square error of cross-validation variation. **(C)** Regression coefficient path map.

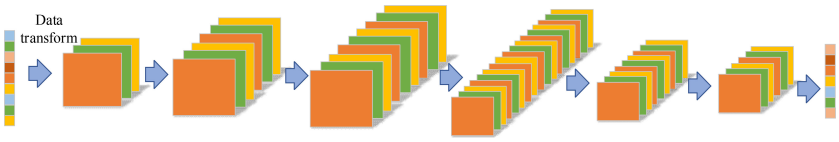


FIGURE 8
Overall network structure of CNN model.

resolutions of 8×8×3, 16×16×3 and 32×32×3. For the prediction part, adaptive maximum pooling operation and Softmax are mainly used to output the prediction results.

3.2 Building and training of model loss function

Because the maize variety classification model was a multi-category model, the Cross Entropy Loss function was used to regress training the maize variety identification model. The specific formula is shown in formula (2). In the formula, y_j represents the unique thermal coding form corresponding to the actual category, and o_j represents the probability that the network predicts a certain category.

$$Loss = \sum_{j=1}^{q} y_j log \sum_{j=1}^{q} \exp(o_j) - \sum_{j=1}^{q} y_j o_j \quad (2)$$

The classification model of maize varieties is constructed with Pytorch framework. The hardware platform is Intel (R) Xeon (R) Silver 4210R, the main frequency is 3.5 GHz, and the memory is 32GB. The graphics card model is NIVIDIA GeForce RTX 2080Ti GPU, and the video memory is 16GB. The software platforms are Pycharm 2020.2, CUDNN 7.4. 1.5, Python 3.8 and Pytorch 1.2. The training period of epoch is set to 10000, and the initial learning rate is set to 0.01. The learning rate (lr) is adjusted periodically by LambdaLR algorithm, and the model parameters are optimized by SGD optimizer in one step to improve the training effect of the model.

3.3 Model evaluation index

To comprehensively evaluate the detection performance of maize variety classification model, training set accuracy (Train), test set accuracy (Test), frames per second (FPS), model weight (Weight), model computation (Flops), model parametric number (Params), Precision and Recall are used as evaluation indexes, and their specific calculation formulas are as follows.

TABLE 1 Parameter index of CNN model.

Fom	Input			Operator	#out	Stride	Layer
0.75	32^2×3	16^2×3	8^2×3	Conv2d 3×3	12	1	1
1.0	32^2×3	16^2×3	8^2×3	Conv2d 3×3	16	1	
1.25	32^2×3	16^2×3	8^2×3	Conv2d 3×3	20	1	
0.75	32^2×12	16^2×12	8^2×12	Conv2d 3×3	24	1	2
1.0	32^2×16	16^2×16	8^2×16	Conv2d 3×3	32	1	
1.25	32^2×20	16^2×20	8^2×20	Conv2d 3×3	40	1	
0.75	32^2×24	16^2×24	8^2×24	Conv2d 3×3	48	2	3
1.0	32^2×32	16^2×32	8^2×32	Conv2d 3×3	64	2	
1.25	32^2×40	16^2×40	8^2×40	Conv2d 3×3	80	2	
0.75	16^2×48	8^2×48	4^2×48	Conv2d 3×3	24	1	4
1.0	16^2×64	8^2×64	4^2×64	Conv2d 3×3	32	1	
1.25	16^2×80	8^2×80	4^2×80	Conv2d 3×3	40	1	
0.75	16^2×24	8^2×24	4^2×24	Conv2d 3×3	12	1	5
1.0	16^2×32	8^2×32	4^2×32	Conv2d 3×3	16	1	
1.25	16^2×40	8^2×40	4^2×40	Conv2d 3×3	20	1	

Fom represents depth factors of different sizes between model layers, Input represents 3D spatial feature matrices of different sizes, Operator represents corresponding convolution operations, out represents the size of feature maps output between model layers, Stride represents the step size of convolution kernel scanning, and Layer represents the names of convolution layers in different stages.

10.3389/fpls.2023.1254548

$$Precision = \frac{TP}{TP + FP} \quad (3)$$

$$Recall = \frac{TP}{TP + FN} \quad (4)$$

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

In the formula, TP represents the number of positive samples that the model prediction is consistent with the real label, FP represents the number of samples that the model prediction is inconsistent with the actual positive samples, FN represents the number of samples that the model prediction is inconsistent with the actual negative samples, and TN represents the number of samples that the model prediction is consistent with the actual negative samples.

4 Results and analysis

In this study, the maize variety identification model was trained and tested by 3-fold cross-validation to verify the applicability of maize variety identification model. When the characteristic dimension is 768, the specific results of cross-validation of three maize variety identification models, $MVI_{0.75}$ D_1, $MVI_{1.0}$ D_1 and $MVI_{1.25}$ D_1, are shown in Table 2. As can be seen from Table 2, the accuracy of training set and test set of $MVI_{1.0}D_1$ are 97.62% and 96.65% respectively, and the performance is the best. Compared with $MVI_{0.75}$ D_1 and $MVI_{1.25}$ D_1, the accuracy of $MVI_{1.0}D_1$ model test set is improved by 7.2% and 1.43%. The inference speeds of $MVI_{0.75}D_1$, $MVI_{1.0}D_1$ and $MVI_{1.25}D_1$ are 666Fps/s, 588Fps/s and 526Fps/s respectively in CPU mode and 1000Fps/s, 1000Fps/s and 909Fps/s in graphics card environment, which shows that the three models have real-time detection performance. Although $MVI_{0.75}D_1$, $MVI_{1.\ 0}D_1$ and $MVI_{1.\ 25}D_1$ are inputted the same spatial feature resolution the model detection results are significantly different in Weight, Params and Flops due to the influence of depth scaling factor of model layers. Weight, Params and Flops of $MVI_{0.75}D_1$ model are 2.33M, 70.35k and 286k respectively, which are the smallest among the three model. Weight, Params and Flops of $MVI_{0.75}D_1$ and $MVI_{1.25}D_1$ models are (4.05M, 6.25M), (90.73k, 116.86k) and (367k, 469k) respectively. Although there are obvious differences in the metrics of $MVI_{0.75}D_1$, $MVI_{1.0}D_1$ and $MVI_{1.25}D_1$, the parameters and calculation amount are still small and negligible compared with the classical CNN model. Therefore, it can be concluded that the detection effect of $MVI_{1.0}D_1$ is the best among $MVI_{0.75}D_1$, $MVI_{1.0}D_1$ and $MVI_{1.25}D_1$ three maize identification models.

To explore the influence of spatial feature dimension on the training results of maize variety identification model, the test results of maize variety identification model with two spatial feature dimensions 192 and 3072 were listed. The specific results are shown in Tables 3 and 4. The comparative analysis shows that the overall performance of the maize variety discrimination model $MVI_{1.25}D_0$ is better when the dimension is192 with the same width scaling factor d. In addition, it can be found from Table 3 that when the width scaling factor d is 1.25, the accuracy rate of maize variety detection model in training set and test set is the best, which can reach 99.20% and 95.34% respectively. This reflects that when the feature space dimension is small, the maize variety identification model searches for effective features less effectively. Therefore, properly adjusting the depth scaling factor is helpful to improve the feature extraction ability and generalization performance of the model. According to the test results in Tables 2–4, it can be found that when the spatial feature resolution is enlarged only by improving the input feature dimension, the performance of the maize variety identification model is not better with a larger input feature dimension. Appropriate adjustment of the spatial feature dimension is helpful to improve the detection effect of the model. The best result is obtained when the dimension is 768, and the accuracy of maize variety identification model is improved obviously. In addition, the reasoning speed of CPU, Weight, Params, and Flops of maize variety identification model increased exponentially when the size of input feature dimension was

TABLE 2 The model results of cross-validation when dimension is 768.

N	Model	Train/%	Test/%	Params	Flops	Weight	Fps_g/s	Fps_c/s
1	$MVI_{0.75}D_1$	91.73	86.12	70.35k	2.33M	286k	0.001	0.0015
2	$MVI_{0.75}D_1$	97.14	89.95	70.35k	2.33M	286k	0.001	0.0015
3	$MVI_{0.75}D_1$	95.55	88.52	70.35k	2.33M	286k	0.001	0.0015
1	$MVI_{1.0}D_1$	97.62	96.65	90.73k	4.05M	367k	0.001	0.0017
2	$MVI_{1.0}D_1$	99.84	95.69	90.73k	4.05M	367k	0.001	0.0017
3	$MVI_{1.0}D_1$	99.5	94.73	90.73k	4.05M	367k	0.001	0.0017
1	$MVI_{1.25}D_1$	98.73	92.82	116.86k	6.25M	469k	0.0011	0.0019
2	$MVI_{1.25}D_1$	99.84	95.22	116.86k	6.25M	469k	0.0011	0.0019
3	$MVI_{1.25}D_1$	99.84	93.78	116.86k	6.25M	469k	0.0011	0.0019

MVI_mD_n represents different classification model of maize varieties. Among them, m represents the depth factor of model layer, m can be taken as 0.75, 1.0 and 1.25, n represents the input feature dimensions of different sizes, and n can be taken as 0, 1 and 2 respectively, which respectively represent the three states that the input feature dimensions are equal to 192, 768 and 3072. Train represents the accuracy of maize variety classification model in training set, and Test represents the accuracy of maize variety classification model in test set. Fps_g represents the frame detection speed in GPU environment, and Fps_c represents the frame detection speed in CPU environment.

10.3389/fpls.2023.1254548

TABLE 3 The model results of cross-validation when dimension is 192.

N	Model	Train/%	Test/%	Params	Flops	Weight	Fps_g/s	Fps_c/s
1	$MVI_{0.75}D_0$	92.69	89.47	37.52k	582.04k	158k	0.001	0.001
2	$MVI_{0.75}D_0$	90.78	85.17	37.52k	582.04k	158k	0.001	0.001
3	$MVI_{0.75}D_0$	89.83	83.73	37.52k	582.04k	158k	0.001	0.001
1	$MVI_{1.0}D_0$	99.68	93.78	57.9k	1.01M	239k	0.001	0.0012
2	$MVI_{1.0}D_0$	94.28	86.60	57.9k	1.01M	239k	0.001	0.0012
3	$MVI_{1.0}D_0$	95.71	90.91	57.9k	1.01M	239k	0.001	0.0012
1	$MVI_{1.25}D_0$	98.57	92.83	84.03k	1.26M	341k	0.0011	0.0013
2	$MVI_{1.25}D_0$	99.20	95.34	84.03k	1.26M	341k	0.0011	0.0013
3	$MVI_{1.25}D_0$	99.52	94.26	84.03k	1.26M	341k	0.0011	0.0013

TABLE 4 The model results of cross-validation when dimension is 3072.

N	Model	Train/%	Test/%	Params	Flops	Weight	Fps_g/s	Fps_c/s
1	$MVI_{0.75}D_2$	81.08	75.60	201.68k	9.31M	800k	0.001	0.0026
2	$MVI_{0.75}D_2$	84.89	79.43	201.68k	9.31M	800k	0.001	0.0026
3	$MVI_{0.75}D_2$	80.45	75.60	201.68k	9.31M	800k	0.001	0.0026
1	$MVI_{1.0}D_2$	88.87	82.78	226.06k	16.29M	880k	0.001	0.0032
2	$MVI_{1.0}D_2$	86.49	83.73	226.06k	16.29M	880k	0.001	0.0032
3	$MVI_{1.0}D_2$	90.62	86.60	226.06k	16.29M	880k	0.001	0.0032
1	$MVI_{1.25}D_2$	98.57	91.39	248.19k	24.99M	982k	0.0011	0.0037
2	$MVI_{1.25}D_2$	99.20	87.56	248.19k	24.99M	982k	0.0011	0.0037
3	$MVI_{1.25}D_2$	99.52	92.34	248.19k	24.99M	982k	0.0011	0.0037

changed, while the reasoning speed of GPU was basically stable at 1000Fps/s.

In order to further explore the influence of input feature dimension and model layer depth on maize variety classification model, the Recall and Precision indexes of nine maize variety classification models were analyzed, as shown in Figure 9. It can be seen from Figure 9 that the Recall and Precision of $MVI_{1.0}D_1$ model are the highest, respectively 96.7% and 96.8%, followed by maize variety classification models with the same characteristic dimension (dimension is768) and different model layer depths. The Recall and Precision of MVI_mD_0 model are more stable than MVI_mD_2, which also proves that the depth of model layer is not positively correlated with the performance of model classification. Appropriate adjustment of model layer depth is helpful to improve the effective extraction of spectral features of maize variety classification network. The variation of loss curves of nine models in 10000 iteration periods is shown in Figure 10. The loss curve of MVI_mD_1 model converges fastest with the increase of iterations and the overall fluctuation is slight. The loss curve of MVI_mD_0 model fluctuates more than that of MVI_mD_1, but the general convergence is faster. The loss curve of MVI_mD_2 is more divergent and the

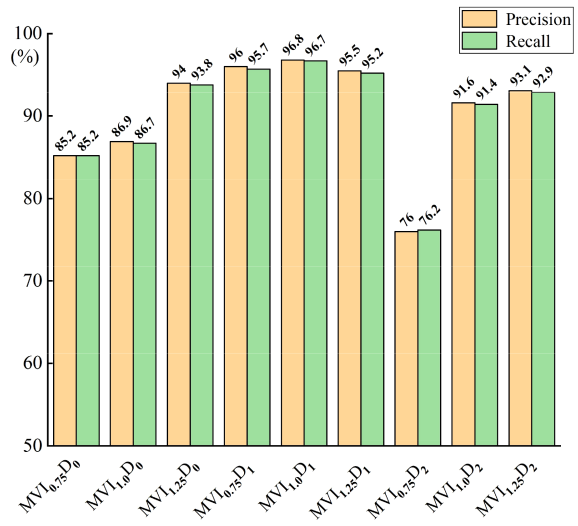


FIGURE 9
Recall and precision results of different models.

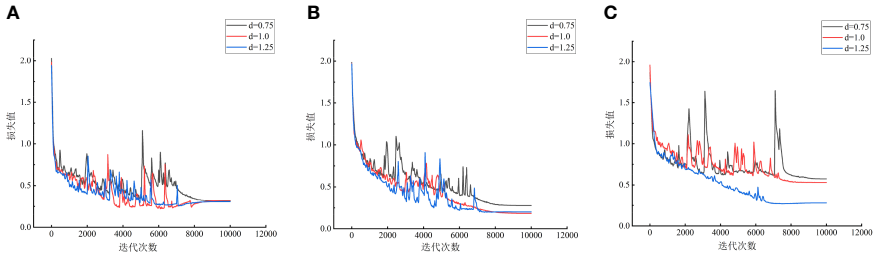


FIGURE 10
Loss curve of different models. **(A)** Dimension is192. **(B)** Dimension is 768. **(C)** Dimension is 3072.

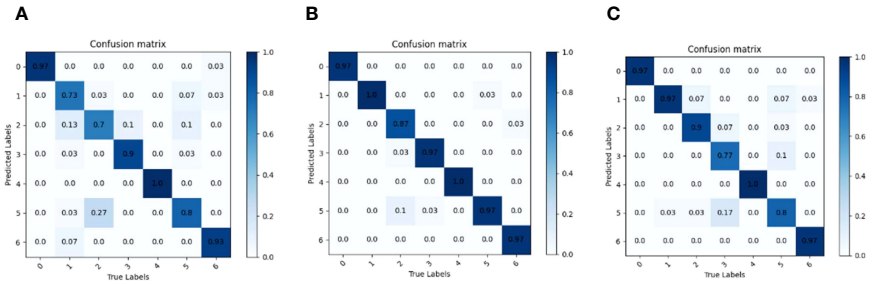


FIGURE 11
Model confusion matrix when layer depth scaling factor d is1. 0. **(A)** $MVI_{1.0}D_0$. **(B)** $MVI_{1.0}D_1$. **(C)** $MVI_{1.0}D_2$.

overall convergence is poorer with the increase of iterations, indicating that when the input dimension is 3072, it is easy to generate invalid feature redundancy, which is not conducive to the extraction of effective features by maize variety classification model.

To explore the effects of MVI_mD_0, MVI_mD_1 and MVI_mD_2 series of maize variety classification models on seven kinds of maize seeds, three maize variety classification models ($MVI_{1.0}D_0$, $MVI_{1.0}D_1$ and $MVI_{1.0}D_2$) with layer depth scaling factor d=1. 0 were selected to test the test set, and the correlation confusion matrix was drawn by comparing the predicted results with the actual results, as shown in Figure 11. It can be seen from Figure 11B that $MVI_{1.0}D_1$ is the best in classifying seven maize seeds and there are no misidentifications and omissions in category 1 and category 4. Although the $MVI_{1.0}D_1$ model shows misrecognition among categories 1, 3, 5 and 6, the misidentification rate is lower compared with the confusion matrix results of $MVI_{1.0}D_0$ and $MVI_{1.0}D_2$, and $MVI_{1.0}D_1$ only misidentifies category 0 without misidentification. Compared with $MVI_{1.0}D_1$, $MVI_{1.0}D_0$ and $MVI_{1.0}D_2$ show more misidentification and the model classifiers are unbalanced.

In this study, seven hybrid maize varieties were taken as the research objects, and the effects of different input feature dimensions and model layer depth on the performance of the maize variety classification model were discussed emphatically. The results showed that the maize variety classification model performs better when the input feature dimension is 768 and worse when the input feature dimension is 3072. This phenomenon may be attributed to the redundant and invalid features easily produced by the higher feature dimension, which indirectly affects the classification effect of the maize variety classification model. Therefore, changing the dimension of input features can effectively improve the ability of extracting effective spectral features of maize variety classification model. In addition, the effect of model layer depth on the performance of maize variety classification model was also discussed in this study. From the results, it can be found that there is no positive correlation between the performance of maize variety classification model and the layer depth of the model. When the layer depth factor d is 1.25, the performance of the maize variety classification model is slightly lower than that when d is 1. 0, so it is most appropriate to set the layer depth factor d as 1. 0. Due to the small sample size, more sample data will be collected in the future to further validate the maize classification model whether the method of identifying maize varieties by mapping characteristic bands to high-dimensional spatial features is feasible.

5 Conclusion

(1) To solve the problem of less effective feature bands and lack of information by single feature variable extraction method, 56 feature bands are selected by combining SPA and CARS in this study.

(2) To solve the problems of poor effect and slow speed of traditional machine learning method in maize classification, a high-dimensional feature mapping method is adopted to

reshape the extracted feature bands into three-dimensional image features after mapping them to a high-dimensional space. And a five-layer convolution neural network is constructed to identify three-dimensional image features.

(3) At the same time, the influence of the size of the input feature dimension and the depth of the model layer on the performance of the maize variety model are discussed in this study. The test results show that when the dimension of the input feature dimension is 768 and the depth factor of the layer is 1.0, the performance of maize variety classification model is the best. And the accuracy of the test set is 96.65%, and the detection frame rate is 1000Fps/s in GPU environment, which can realize rapid and effective non-destructive detection of maize varieties.

Data availability statement

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

Author contributions

FZ: Funding acquisition, Investigation, Project administration, Resources, Supervision, Writing – review & editing. FYZ: Data curation, Formal Analysis, Writing – original draft, Writing –review & editing. SQW: Formal Analysis, Methodology, Software, Supervision, Writing – original draft. LTL: Investigation, Resources, Supervision, Writing –review & editing. QL: Data curation, Funding acquisition, Investigation, Resources, Supervision, Writing – review & editing. SLF: Investigation, Resources, Supervision, Writing – review & editing. XYW: Supervision, Writing – review & editing. QFL: Investigation, Resources, Writing – review & editing. YKZ: Investigation, Supervision, Writing – review & editing.

Funding

This Paper is supported by the National Natural Science Foundation of China (Grant No. 52075149), Frontier Exploration Projects of Longmen Laboratory (Grant No. LMQYTSKT032), the National Key Research and Development Program of China project (No. 2017YFD0301106), the Scientific and Technological Project of Henan Province (Grant No. 232102111119 and 222102110196), Colleges and Universities of Henan Province Youth Backbone Teacher Training Program (Grant No. 2017GGJS062).

Conflict of interest

Author QL was employed by the company Henan Pingan Seed Industry Limited Company.

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

Publisher's note

References

Chivasa, W., Mutanga, O., and Biradar, C. (2019). Phenology-based discrimination of maize (Zea mays L.) varieties using multitemporal hyperspectral data. *J. Appl. Remote Sens.* 13, 1. doi: 10.1117/1.JRS.13.017504

Feng, R., Chen, Z., and Yi, S. (2022). Identification of corn varieties based on bayesian optimization SVM. *Spectrosc. Spectral. Analysis.* 42 (06), 1698–1703. doi: 10.3964/j.issn.1000-0593(2022)06-1698-06

Feng, Z., Zhu, Q., and Huang, M. (2012). Maize variety recognition using hyperspectral image. *J. Jiangnan. Univ. Natural Sci. Edition.* 11 (02), 149–153. doi: 10.3969/j.issn.1671-7147.2012.02.005

Fu, L., Sun, J., Wang, S., Xu, M., Yao, K., Cao, Y., et al. (2022). Identification of maize seed varieties based on stacked sparse autoencoder and near-infrared hyperspectral imaging technology. *J. Food Process. Eng.* 45, 1–12. doi: 10.1111/jfpe.14120

Huang, M., He, C., Zhu, Q., and Qin, J.. (2016a). Maize seed variety classification using the integration of spectral and image features combined with feature transformation based on hyperspectral imaging. *Appl. Sci.* 6)6, 183. doi: 10.3390/app6060183

Huang, J., Liu, C., Song, S., Tan, Y., Han, Y., Xu, L., et al. (2022). Research progress of spectroscopic techniques in quality detection of maize varieties. *Agric. Technol.* 42 (02), 5–8. doi: 10.19754/j.nyyjs.20220130002

Huang, M., Tang, J., et al. (2016b). Classification of maize seeds of different years based on hyperspectral imaging and model updating. *Comput. Electron. Agric.* 122, 139–145. doi: 10.1016/j.compag.2016.01.029

Shao, Q., Chen, Y., Yang, S., Zhao, Y., and Li, J. (2019). Identification of maize seed varieties based on random forest and hyperspectral technique. *Geogr. Geo-Information. Sci.* 35 (05), 34–39. doi: 10.3969/j.issn.1672-0504.2019.05.006

Sun, M., Huang, Y., Chen, X., Liu, Y., and Su, Q. (2022). Rapid research on the storage year of wheat seed based on hyperspectrum imaging. *J. Chin. Cereals. Oils. Assoc.* 37 (01), 170–174. doi: 10.3969/j.issn.1003-0174.2022.01.026

Sun, Y., Zhang, Y., Shen, F., Li, G., Xing, C., and Yuan, J.. (2022). Identification method of wheat grain mildew based on hyperspectral imaging technology. *J. Chin. Cereals. Oils. Assoc.* 37 (09), 40–46. doi: 10.3969/j.issn.1003-0174.2022.09.008

Tu, K., Wen, S., Cheng, Y., Hu, Y., Pan, T., Hou, H., et al. (2021). A non-destructive and highly efficient model for detecting the genuineness of maize variety 'JINGKE 968'using machine vision combined with deep learning. *Comput. Electron. Agric.* 182, 106002. doi: 10.1016/j.compag.2021.106002

Tu, K., Wen, S., Cheng, Y., Hu, Y., Pan, T., Hou, H., et al. (2022). A model for genuineness detection in genetically and phenotypically similar maize variety seeds based on hyperspectral imaging and machine learning. *Plant Methods* 18, 1–17. doi: 10.1186/s13007-022-00918-7

Wang, W., and Guo, L. (2019). Effects of storage temperature on physiological changes of corn germ and endosperm. *Cereals. Oils.* 32 (05), 68–71. doi: 10.3969/j.issn.1008-9578.2019.05.018

Wang, Z., Huang, W., Tian, X., Long, Y., Li, L., and Fan, S. (2022). Rapid and non-destructive classification of new and aged maize seeds using hyperspectral image and chemometric methods. *Front. Plant Sci.* 13, 1–12. doi: 10.3389/fpls.2022.849495

Wang, L.-G., and Wang, L.-F. (2021). Variety identification model for maize seeds using hyperspectral pixel-level information combined with convolutional neural network. *Natl. Remote Sens. Bull.* 25 (11), 2234–2244. doi: 10.11834/jrs.20219349

Wang, D., Wang, K., Wu, J., and Han, P. (2021). Progress in research on rapid and non-destructive detection of seed quality based on spectroscopy and imaging technology. *Spectrosc. Spectral. Analysis.* 41 (01), 52–59. doi: 10.3964/j.issn.1000-0593(2021)01-0052-08

Wu, Y., Li, M., Zhang, B., Zhang, Y., and Guo, B. (2021). Application of hyperspectral imaging technology in grain quality detection. *J. Chin. Cereals. Oils. Assoc.* 36 (05), 165–173. doi: 10.3969/j.issn.1003-0174.2021.05.027

Wu, X., Zhang, W., Lu, J., Qiu, Z., and He, Y. (2016). Study on visual identification of corn seeds based on hyperspectral imaging technology. *Spectrosc. Spectral. Analysis.* 36 (02), 511–514. doi: 10.3964/j.issn.1000-0593(2016)02-0511-04

Wu, J., Zhang, L., Li, J., Liu, C., Sun, X., and Yu, L. (2022). Detection model of moisture content of single maize seed based on hyperspectral image and ensemble learning. *Trans. Chin. Soc. Agric. Machinery.* 53 (05), 302–308. doi: 10.6041/j.issn.1000-1298.2022.05.031

Yang, B., Duan, M., and Yang, T. (2022). Research on the classification method of watermelon seed vigor level based on hyperspectral imaging technology. *J. Henan. Agric. Sci.* 51 (09), 151–158. doi: 10.15933/j.cnki.1004-3268.2022.09.016

Zhang, H., Yao, C., Jiang, M., Ji, Y., and Li, H. (2019). Research on wheat seed classification and recognition based on hyperspectral imaging. *J. Triticeae. Crops* 39 (01), 96–104. doi: 10.7606/j.issn.1009-1041.2019.01.13

frontiers | Frontiers in Plant Science

TYPE Original Research
PUBLISHED 11 October 2023
DOI 10.3389/fpls.2023.1257212

OPEN ACCESS

EDITED BY
Xi Tian,
Beijing Academy of Agriculture and Forestry Sciences, China

REVIEWED BY
Guoxiong Zhou,
Central South University Forestry and Technology, China
Aibin Chen,
Central South University Forestry and Technology, China

*CORRESPONDENCE
Yange Sun
yangesun@xynu.edu.cn

RECEIVED 12 July 2023
ACCEPTED 19 September 2023
PUBLISHED 11 October 2023

CITATION
Sun Y, Wu F, Guo H, Li R, Yao J and Shen J (2023) TeaDiseaseNet: multi-scale self-attentive tea disease detection.
Front. Plant Sci. 14:1257212.
doi: 10.3389/fpls.2023.1257212

TeaDiseaseNet: multi-scale self-attentive tea disease detection

Yange Sun[1,2*], Fei Wu[1], Huaping Guo[1,2], Ran Li[1], Jianfeng Yao[1,3] and Jianbo Shen[4]

[1]School of Computer and Information Technology, Xinyang Normal University, Xinyang, China, [2]Research Center of Precision Sensing and Control, Institute of Automation, Chinese Academy of Sciences, Beijing, China, [3]Henan Key Laboratory of Tea Plant Biology, Xinyang Normal University, Xinyang, China, [4]Intelligent Equipment Research Center, Beijing Academy of Agriculture and Forestry Sciences, Beijing, China

Accurate detection of tea diseases is essential for optimizing tea yield and quality, improving production, and minimizing economic losses. In this paper, we introduce TeaDiseaseNet, a novel disease detection method designed to address the challenges in tea disease detection, such as variability in disease scales and dense, obscuring disease patterns. TeaDiseaseNet utilizes a multi-scale self-attention mechanism to enhance disease detection performance. Specifically, it incorporates a CNN-based module for extracting features at multiple scales, effectively capturing localized information such as texture and edges. This approach enables a comprehensive representation of tea images. Additionally, a self-attention module captures global dependencies among pixels, facilitating effective interaction between global information and local features. Furthermore, we integrate a channel attention mechanism, which selectively weighs and combines the multi-scale features, eliminating redundant information and enabling precise localization and recognition of tea disease information across diverse scales and complex backgrounds. Extensive comparative experiments and ablation studies validate the effectiveness of the proposed method, demonstrating superior detection results in scenarios characterized by complex backgrounds and varying disease scales. The presented method provides valuable insights for intelligent tea disease diagnosis, with significant potential for improving tea disease management and production.

KEYWORDS

tea disease detection, deep learning, multi-scale feature, self-attention, convolutional neural networks

1 Introduction

As one of the traditional cash crops, tea holds significant economic and cultural value. However, the susceptibility of tea plants to diseases during their growth stages has a detrimental effect on both yield and quality, leading to significant economic losses for tea growers (Mukhopadhyay et al., 2021; Mahum et al., 2023; Sunil et al., 2023). Conventional manual techniques for detecting tea diseases are laborious, time-consuming, and

dependent on the expertise of the testers, leading to inefficiency and high expenses (Drew, 2019; Abade et al., 2021). Additionally, the complex tea plantation environment, including elements like weeds, branches, and soil, along with factors like varying disease scales and densely shaded areas, pose challenges for accurately detecting of tea diseases. Therefore, there is an urgent need for research on rapid and precise methods for early detection of tea diseases. Implementing such methods would enable tea farmers to promptly implement control measures, prevent disease spread, protect the health of tea plantations, and promote the sustainable development of the tea industry (Debnath et al., 2021; Lanjewar and Panchbhai, 2023; Pandian et al., 2023).

Traditional machine learning models, such as support vector machines (SVM), decision trees, K-means, and random forests, require manual feature design specific to different disease types, making them susceptible to environmental factors and lacking generalization capabilities (Bhavsar et al., 2022; Zou et al., 2020; Steven, 2021; Yu et al., 2021; Bao et al., 2022; Prabu et al., 2022). Conversely, deep learning, particularly in object detection, exhibits potential in crop disease identification (Krisnandi et al., 2019; Ayan et al., 2020; Jiang et al., 2020; Tetila et al., 2020; Xiong et al., 2020; Hu et al., 2021b). However, existing models that solely consider local pixel relationships at short distances struggle to incorporate crucial global information in complex scenarios of tea disease detection, featuring varying disease scales and complex backgrounds, leading to limitations in detection accuracy (Li et al., 2021).

Convolutional Neural Networks (CNNs) have demonstrated remarkable success in automatically learning multi-level, high-order features from disease images, surpassing the limitations of traditional manual feature design methods (Abade et al., 2021; Akanksha et al., 2021; Dhaka et al., 2021; Latha et al., 2021; Lu et al., 2021; Wang et al., 2021; Yogeshwari and Thailambal, 2021; Ashwinkumar et al., 2022). They offer significant advantages in disease detection and have been extensively studied (Liu et al., 2022; Kirti and Rajpal, 2023; Kirti et al., 2023; Sudhesh et al., 2023; Tholkapiyan et al., 2023; Xu et al., 2023; Zhou et al., 2023). Depending on their network structure, CNN-based disease detection methods can be categorized as one-stage or two-stage detectors (Jiao et al., 2021; Lin et al., 2023). Regarding tea disease detection techniques, Qi et al. introduced TC-YOLO, a lightweight deep learning architecture based on YOLO that achieves high fusion capabilities (Qi et al., 2022). Alruwaili et al. improved the Faster R-CNN model for disease detection and achieved better recognition performance than other models (Alruwaili et al., 2022). By utilizing basic convolutional layer architectures, Lee et al. achieved an accuracy of 77.5% in detecting insect pests and diseases (Lee et al., 2020). Hu et al. introduced an algorithm that enhances image quality to improve detection accuracy (Hu et al., 2021a). Chen et al. developed LeafNet, a specialized CNN model for tea disease feature extraction (Chen et al., 2019). Xue et al. proposed YOLO-tea, a tea disease detection model based on YOLOv5 (Xue et al., 2023). However, CNNs overlook crucial global information among distant pixels, which impacts detection accuracy. Researchers are currently exploring methods to enhance the global modeling capabilities of CNNs in these scenarios. For instance, Hou et al. proposed an improved two-stage Faster R-CNN disease detection algorithm incorporating an attention mechanism in the network (Hou et al., 2023).

Attention mechanisms have emerged as highly successful approaches in disease detection tasks, aiming to emulate the remarkable capabilities of the human visual system in capturing vital information from complex scenes (Zheng et al., 2021; Hu et al., 2023). Spatial attention, channel attention, and self-attention are different attention mechanisms that enhance feature extraction and model performance (Carion et al., 2020; Guo et al., 2022). Several studies have employed attention mechanisms in disease detection models. For instance, Liu et al. proposed the spatial attention module (Liu et al., 2019), Wang et al. introduced both channel and spatial attention mechanisms (Wang et al., 2020), Zha et al. developed a lightweight network model based on a coordinate attention mechanism (Zha et al., 2021), Zhu et al. combined CNNs with Transformer architecture to establish (Zhu et al., 2022). Similarly, Lin et al. proposed a YOLO-based algorithm that employs a self-attentive mechanism to enhance the model's ability to capture global information on tea diseases (Lin et al., 2023). Borhani et al. proposed combining CNNs with Transformer architecture to exploit the Transformer's capability to establish dependencies between distant features and extract global disease features (Borhani et al., 2022). By incorporating attention mechanisms, researchers have made considerable progress in capturing essential information and enhancing the performance of disease detection models (Alirezazadeh et al., 2023; Yang et al., 2023).

Although the studies mentioned above have made progress in considering local disease information, it is crucial to emphasize the value of global information, especially the interaction between distant pixels (Sapna et al., 2023). Motivated by these challenges and research gaps, we introduce a novel network named Tea Disease Network (TeaDiseaseNet). Our proposed network integrates multi-scale feature representation with a self-attention mechanism to enhance performance in complex backgrounds and variable disease scales. The primary contributions can be summarized as follows:

(1) Introducing the Multi-scale Feature Extraction Module (MFEM), which utilizes multi-scale convolutional neural networks (CNNs) to capture comprehensive and localized multi-scale feature representations from disease images effectively. This module facilitates the extraction of comprehensive local spatial information.

(2) Devising the Scale Self-Attentive Module (SSAM) to address scale variations and complex backgrounds. The SSAM incorporates self-attention blocks to consolidate local and global information on tea disease images, facilitating effective interaction between global information and local features.

(3) Designing the Scale-aware Feature Fusion module (SFF) to achieve accurate and robust detection. The SFF enables feature fusion and the network to distinguish the relative importance of different input features. It enhances the accuracy and robustness of tea disease detection by facilitating multi-scale feature fusion.

(4) Conducting extensive comparative experiments and ablation studies on each module to demonstrate our proposed method's superior performance and effectiveness. The results show significant improvements in various scenarios, surpassing most existing methods. These findings highlight the potential and effectiveness of our approach in enhancing the detection of tea diseases.

The structure of this paper is as follows. Section 2 focuses on the dataset utilized in this research and explains the enhanced modules integrated into TeaDiseaseNet. Section 3 covers the experimental setup, including equipment configuration, evaluation criteria, and experimental parameters. We present the results and analysis of the ablation experiments, visualization, and discussion. Finally, in Section 4, we present our conclusions and discuss potential avenues for future research.

2 Materials and methods

In this section, we outline the critical components of our proposed TeaDiseaseNet detection method. Our method involves two main aspects: collecting a comprehensive tea disease dataset and developing of an accurate disease detection framework. The dataset collection process includes acquiring disease images, annotating the dataset, and appropriately partitioning it. The detection model comprises three crucial functional modules: 1) The Multi-scale Feature Extraction Module (MFEM) extracts features from different scales to capture detailed information about tea diseases. 2) The Scale Self-Attention Module (SSAM) applies self-attention mechanisms to learn contextual dependencies within the extracted features. 3) The Scale-aware Feature Fusion (SFF) module fuses the multi-scale and self-attended features to generate a robust representation for disease detection. Collectively, these components contribute to the effectiveness and accuracy of our TeaDiseaseNet detection.

2.1 Tea disease dataset construction

2.1.1 Disease images acquisition

The tea disease dataset utilized in the experiments was obtained from Professor Jiang Zhaohui's research group at Anhui Agricultural University (Tholkapiyan et al., 2023). This dataset consists of 776 samples and covers a wide range of tea diseases, including tea exobasidium blight, tea red scab, tea cloud leaf blight, tea cake, tea red rust, and tea algae leaf spot. Each sample image in the dataset has a resolution of 906×600 pixels, ensuring a clear and detailed representation of the tea diseases.

Incorporating diverse tea diseases into the dataset enables comprehensive training and evaluation of the proposed detection model. By including samples from different tea diseases, the dataset offers a rich and representative collection of real-world scenarios encountered by tea growers.

Figure 1 visualizes the dataset, displaying selected tea images that exemplify instances of the six tea diseases above. These images serve as valuable references for understanding each tea disease visual characteristics and distinguishing features. The annotated dataset ensures accurate labeling and facilitates the development of an effective convolutional neural network for tea disease detection.

By utilizing this meticulously collected and annotated dataset, we aim to construct a robust and reliable detection model capable of accurately identifying and classifying tea diseases. The dataset serves as a crucial foundation for our research, ensuring the validity and effectiveness of our proposed TeaDiseaseNet.

2.1.2 Data labeling

To adapt the dataset for tea disease detection tasks, we enhanced the original samples by manually annotating the bounding boxes of the tea disease targets. This critical step involved meticulously labeling each sample image to indicate the precise location and extent of the tea disease instances. The annotation process was performed by a skilled professional using the labelimg image labeling tool, ensuring accuracy and consistency throughout the dataset.

By providing bounding box annotations, we enable the tea disease detection model to identify the presence of tea diseases localize and delineate the specific areas affected by the diseases. This level of detail enhances the model's ability to provide valuable insights and facilitate targeted intervention strategies for tea growers.

The inclusion of bounding box annotations in the dataset enhances its suitability and efficacy for tea disease detection tasks. When used with our advanced TeaDiseaseNet algorithm, the annotated dataset enables accurate and precise detection of tea diseases.

2.1.3 Data augmentation and division

To enhance the model's generalization capability and improve its performance in real-world scenarios, data augmentation techniques were applied to augment the tea disease dataset, thereby expanding its size and diversity. Various methods introduced diversity and variability into the original images, including 90-degree clockwise and counterclockwise rotations, random rotation, noise addition, and exposure adjustments. As a result, a total of 7 640 augmented samples were generated, enriching the dataset and providing a more comprehensive range of training examples for the model.

The augmented dataset was subsequently divided into an 8:2 ratio for training and validation purposes. This division ensured a balanced distribution of data and enabled robust model evaluation. By training the model on a diverse augmented dataset and validating it on separate samples, we obtained more reliable and accurate results. The use of data augmentation techniques, along with the appropriate dataset division, enhances the model's ability to accurately detect tea diseases, even when faced with previously unseen or challenging images.

2.1.4 Characteristics of disease dataset

The dataset's statistical analysis and ranking of scales revealed a significant range of sizes among the tea disease targets. Around 20% of the targets exhibited scales smaller than 0.0207, while 34% had

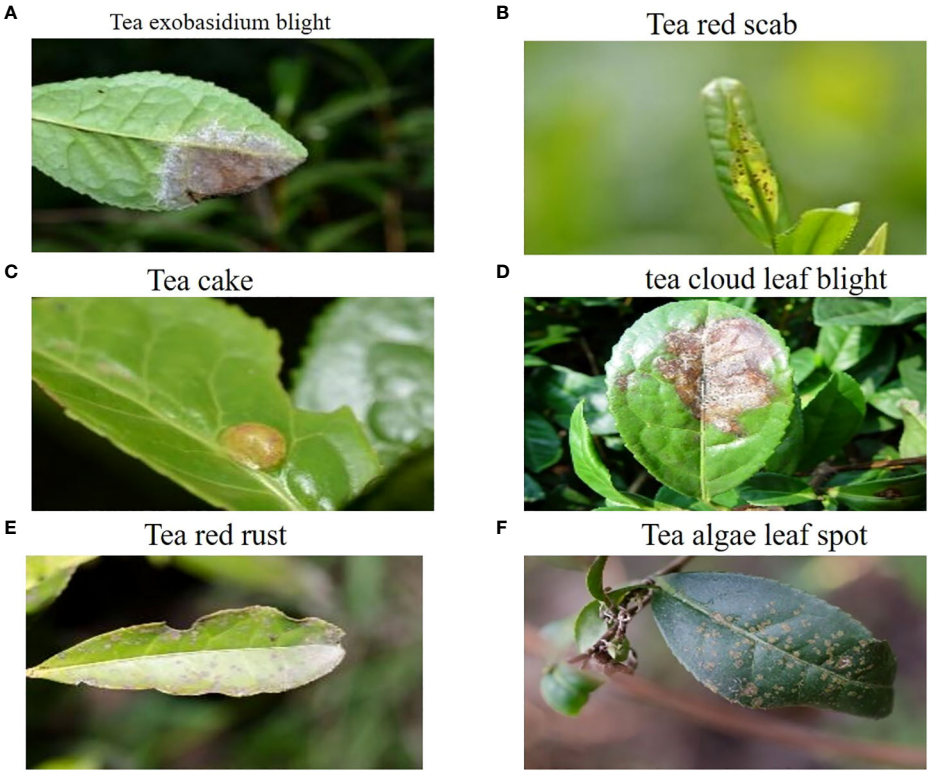


FIGURE 1
Representative Samples from the Tea Dataset.

scales larger than 0.345. This wide range of scales underlines the diverse nature of the dataset and emphasizes the challenge of accurately detecting diseases across various sizes. Understanding these scale variations is crucial for developing a robust detection model capable of effectively identifying tea diseases, regardless of their size. Our goal is to enhance the performance and reliability of the model in detecting tea diseases by addressing the scale variations.

2.2 The architecture of TeaDiseaseNet

To address the challenges posed by variable scales of tea pests and dense, obscuring diseases, this paper presents a novel fused multi-scale self-attentive tea disease detection network based on improving YOLOv5 (Jocher et al., 2022). The YOLOv5 framework is well-known for its remarkable object detection capabilities and high efficiency. In our proposed model, we have harnessed the advantages of YOLOv5 by incorporating multi-scale convolution and multi-scale self-attention mechanisms to effectively capture both local and global features in tea disease images. Figure 2 illustrates the network structure of our model, which comprises three key modules: the Multi-scale Feature Extraction Module (MFEM), the Scale-Self-Attention Module (SSAM), and the Scale-aware Feature Fusion (SFF). These modules synergistically work to achieve accurate and robust tea disease detection. Our approach involves the following steps:

Step 1: Multi-scale feature extraction

We utilize the multi-scale convolutional blocks of the MFEM as a backbone network to extract features from tea images. This allows us to capture feature information of tea diseases at different scales and local levels.

Step 2: Scale self-attentive mechanism

We feed the multi-scale feature maps into the SSAM simultaneously to enable the interaction of global and local information. This mechanism dynamically adjusts the weights of each scale feature, improving the model's ability to capture the characteristics of tea diseases.

Step 3: Scale-aware feature fusion

We incorporate a channel attention mechanism to perform a weighted fusion of features at different scales in tea leaf images. This mechanism efficiently integrates characteristic information of tea diseases across a wide range of scales, enhancing the precision of disease localization and recognition.

Step 4: Prediction

The prediction module utilizes the previously extracted feature information to efficiently localize and identify tea disease features in complex contexts and at varying scales.

By following these steps, our approach aims to effectively extract and integrate features to accurately detect and recognize tea diseases.

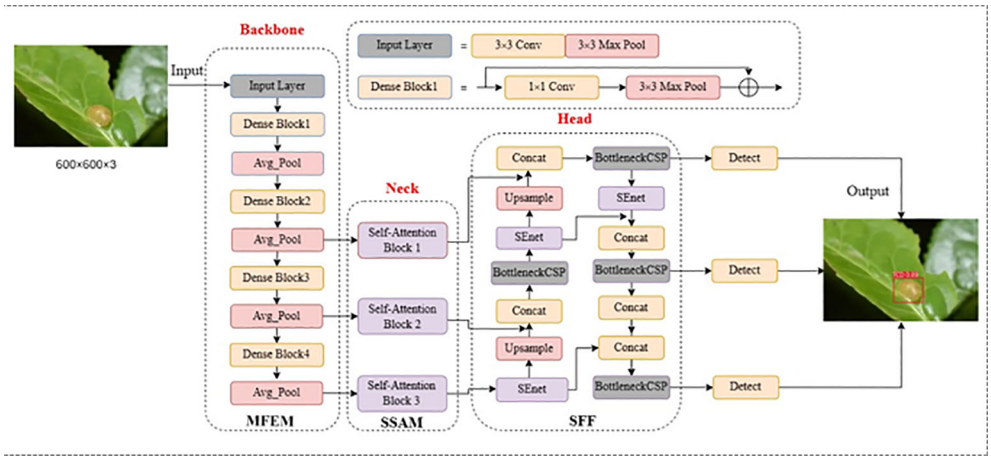


FIGURE 2
The framework of TeaDiseaseNet.

2.3 Multi-scale feature extraction module

Traditional image feature extraction methods often focus on either local or global information, limiting their ability to comprehensively capture the diversity and complexity of images. In recent years, deep learning-based approaches, particularly Vision Transformers (ViT) (Khan et al., 2022), have become the dominant method for image feature extraction. ViT segments images into patches or tokens and employs self-attentive mechanisms to extract parameterized visual representations. However, these methods are constrained by fixed-scale token sequences, which restrict their ability to capture feature structures across different scales. This limitation poses a challenge in tea disease detection due to scale variations. Moreover, self-attentive mechanisms prioritize global information, disregarding important local feature details and blurring the distinction between intricate backgrounds and foregrounds in tea disease images. Consequently, their applicability in disease feature extraction tasks is limited.

To address these challenges, we propose two solutions. The first solution, illustrated in Figure 3A, involves constructing serial multi-scale token sequences by up/down sampling and expanding/reducing token sequences within the self-attentive mechanism module. The second solution, depicted in Figure 3B, consists of constructing parallel multi-scale token sequences wherein images of different scales are simultaneously fed into the self-attentive mechanism module. This approach leverages multi-headed self-attention to capture global contextual information across diverse scales. Compared to the first solution, the second approach provides a simpler implementation. Building on these observations, we propose a parallel multi-scale tea disease feature extraction module to address the limitations of limited local feature representation and a single scale.

As illustrated in Figure 2, our proposed method employs four Dense blocks derived from DenseNet (Roy and Bhaduri, 2022) to extract both the multi-scale structure and local features of tea disease images. The tea disease image passes through the input layer, further progressing into the dense block, and finally undergoing average pooling. The shallow convolutional layers in this module aim to capture intricate features like edges and contours, while the deeper convolutional layers encode comprehensive semantic information. Each level of the Dense block includes down-sampling operations, gradually reducing the resolution of the disease images. We generate a multi-scale feature map by preserving the outputs of the last three levels of Dense blocks. By employing the scaled feature map sequence obtained from the CNN as input for the self-attentive mechanism module, the length of the token sequence is indirectly adjusted. This modification enables each token to represent a larger region in the original image, encompassing a broader range of spatially localized information.

In summary, the MFEM module retrieves multi-scale features, allowing the model to capture information at different levels of detail. This capability is advantageous for tea disease detection tasks as it effectively handles disease size, location, and complex backgrounds variations. The refined multi-scale features enhance the reliability and accuracy of the tea disease detection model.

2.4 Scale self-attention module

The SSAM enables the interaction and fusion of feature maps at various scales using the self-attention mechanism. This allows the tea disease detection model to effectively capture both global and local information in disease images. More specifically, the self-attentive block within the SSAM module takes in multi-scale feature maps as inputs, with each scale's feature maps obtained through convolution. By enhancing information fusion and interaction, this module significantly improves the model's performance and accuracy across various scales. The self-attention operation in each head of the multi-head attention mechanism is computed based on Equation (1).

$$A_{Attention}(Q,K,V) = S_{softMax}\left(\frac{QK^T}{\sqrt{d}} + B\right)V \quad (1)$$

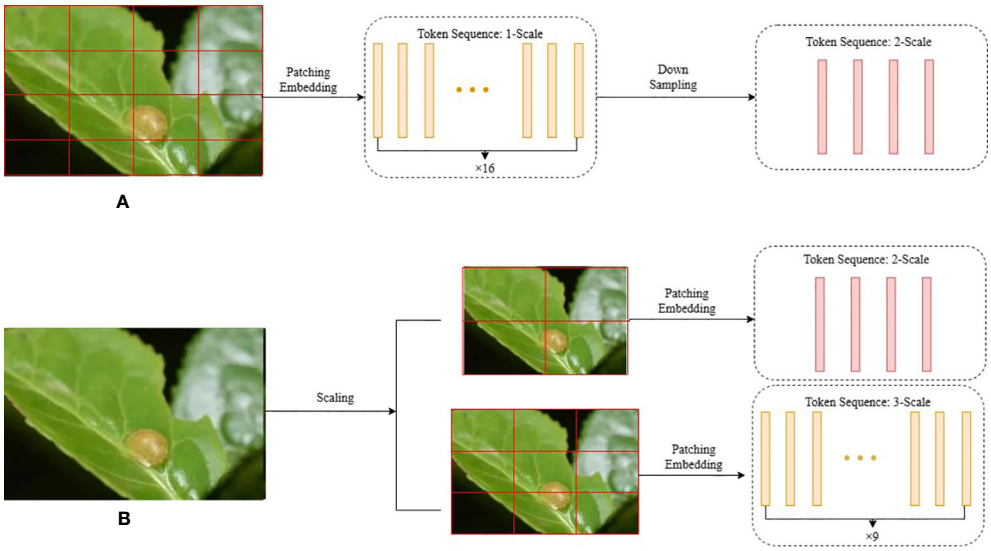


FIGURE 3
Two Ideas for Multi-scale Structures in Tea Disease Detection. **(A)** Constructing serial multi-scale token sequences. **(B)** Constructing parallel multi-scale sequences.

where Q, K, and V represent the query, key, and value matrices, respectively. The vector dimension is denoted as d, and B signifies the bias matrix. The output is obtained by applying the softmax activation function S_{softMax} for multi-classification.

In particular, the Self-Attention Block within the SSAM takes multi-scale feature maps as input. Each scale is obtained through a convolution operation. The configuration of the Self-Attention Block, illustrated in Figure 4, includes a Multi-head Self-Attention (MSA) module that employs a window-based approach and a 2-layer Multi-layer Perceptron (MLP) module. Layer Normalization (LN) layers are incorporated before each MSA and MLP module, and residual connections are employed after each module. This arrangement facilitates the calculation of output features, as shown in Equation (2).

$$\begin{cases} \hat{z}^i = F_{W-MSA}\left(F_{LN}\left(z^{i-1}\right)\right) + z^{i-1} \\ z^i = F_{MLP}\left(F_{LN}\left(\hat{z}^i\right)\right) + \hat{z}^i \\ \hat{z}^{i+1} = F_{SW-MSA}\left(F_{LN}\left(z^i\right)\right) + z^i \\ z^{i+1} = F_{MLP}\left(F_{LN}\left(\hat{z}^{i+1}\right)\right) + \hat{z}^{i+1} \end{cases} \tag{2}$$

where W-MSA represents the window multi-head self-attention, while SW-MSA denotes the shifted window multi-head self-attention. The variables $\hat{z}^i$ and z^i correspond to the output features of the (S)W-MSA and MLP modules of the i-th block, respectively. The W-MSA module, SW-MSA module, MLP module and LN module features are denoted as $F_{W\text{-}MSA}$, $F_{SW\text{-}MSA}$, F_{MLP}, F_{LN}, respectively.

2.5 Scale-aware feature fusion

The SFF module efficiently combines features from multiple scales, resolving discrepancies and improving model performance. In tea disease detection tasks, it is crucial to efficiently process information from multiple scales. This module is specifically designed to address discrepancies and inconsistencies in multi-scale features. We leverage a channel focus mechanism to enhance the model's performance by incorporating spatial and channel features in the input data. This allows the model to accurately discern and differentiate between objects or features, improving object detection accuracy.

The channel attention mechanism enhances the inter-channel information importance in a convolutional neural network. It compresses the features of each channel by integrating their spatial information and computes them using global average pooling, as defined below:

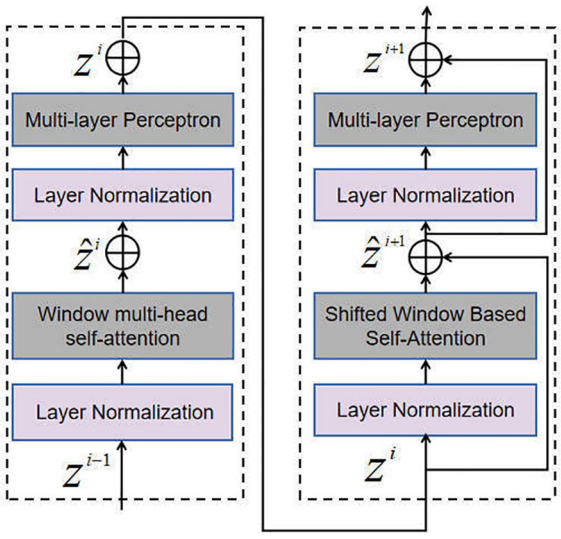


FIGURE 4
Self-Attention Block.

$$z = F_{sq} = \frac{1}{H \times W} \sum_{i=1}^{H} \sum_{i=1}^{W} X(i,j) \tag{3}$$

where z denotes the compressed feature vector, H and W denote the feature map size of feature X. A learnable parameter w captures the correlation between feature channels. To improve computational efficiency, the number of channels is reduced using the following approach:

$$s = F_{ex}(z, w) = \partial(g(z, w)) = (w_2 \delta(w_1 z)) \tag{4}$$

where the adaptive weight of each channel is represented by s, and δ represents the ReLU activation function, while σ represents the Sigmoid activation function. Combining the channel adaptive weight s with the original feature z and assigning a new adaptive weight to each existing channel, the rescaled feature is obtained using Equation (5).

$$Xc = F_{scale}(X, s) = X \cdot s \tag{5}$$

As shown in Figure 5, the SFF consists of Upsample, Concat, Bottleneck CSP, and S_t module operations. The BottleneckCSP module performs a convolution operation on the fused features to further extract feature information, and the St module introduces a channel attention mechanism to weigh the multi-scale features for fusion and eliminate redundant information.

The S_t module utilizes global average pooling to compute feature compression values for each channel and learns parameters to model the correlation between channels, resulting in adaptive weights. These weights are applied to rescaled original features, achieving adaptive feature weighting and improving feature representation. Through the combined operations of Upsample, Concat, BottleneckCSP, and S_t, the feature fusion network enables the interaction and fusion of multi-scale information, enhancing the model's performance. This addresses inconsistencies and discrepancies between multi-scale features, improving accuracy and robustness in tea disease detection tasks.

In general, the primary objective of the SFF module is to integrate global and local information from multiple scales, enabling the generation of precise density maps to effectively capture the spatial distribution of diseases.

2.6 Prediction module

The prediction module is responsible for locating and identifying tea disease information at various scales. It achieves efficient prediction by utilizing the feature fusion network's output and employing parallel branches. The incorporation of IoU branches further enhances the accuracy of the prediction results.

The prediction module comprises 1×1 convolutional layers and parallel branches. Each branch contains a Conv Block that comprises two 3×3 convolutions. The topmost Conv Block is dedicated to the classification task, while the bottommost Conv Block focuses on the regression task. An additional IoU branch is introduced to the module to enhance the accuracy of the predictions.

2.7 Loss functions

The tea disease detection model utilizes three essential loss functions: localization loss L_{loc}, classification loss L_{cls}, and confidence loss L_{conf}, as depicted in Equation (6).

$$L = L_{loc} + L_{cls} + L_{conf} \tag{6}$$

These components enhance the model's accuracy regarding object localization and category identification. The localization loss minimizes bounding box coordinate discrepancies, while the classification loss reduces errors in tea disease classification. Finally, the confidence loss enhances the model's precise detection and classification estimation. The model can optimize its performance by incorporating these loss functions and achieve more accurate and reliable tea disease detection results.

The final localization loss L_{loc} is computed according to Equation (7).

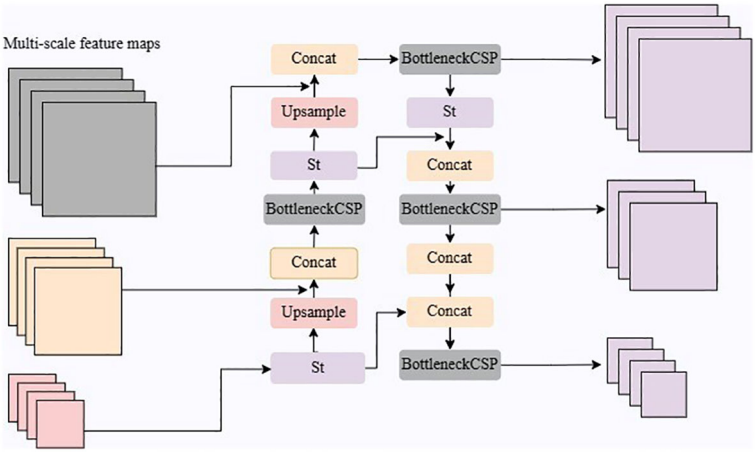


FIGURE 5
The multi-scale feature fusion module.

$$L_{loc} = \sum_{i=1}^{s^2}\sum_{j=0}^{B} I_{i,j}^{obj}(1 - CIoU) \tag{7}$$

The Complete Intersection over Union (CIoU) loss is a regression loss function considering bounding boxes' overlapping area, center distance, and aspect ratio consistency. When incorporated into the model, it provides a more accurate measure of the bounding box regression error, leading to improved accuracy and localization performance in tea disease detection.

$$CIoU = IoU - \frac{\rho^2(b, b^{gt})}{c^2} - \alpha v \tag{8}$$

$$v = \frac{4}{\pi^2}\left(\arctan\frac{w^{gt}}{h^{gt}} - \arctan\frac{w}{h}\right)^2$$

$$\alpha = \frac{v}{(1 - IoU) + v}$$

where c represents the diagonal distance between the prediction frame and the minimum enclosing area of the ground truth frame, ρ denotes the Euclidean distance function, while b and b^{gt} correspond to the centroids of the prediction frame and the actual frame, respectively. The variable ν indicates the similarity of the aspect ratio, and α is the weighting factor. Additionally, w and h denote the width and height of the prediction frame, respectively.

A binary cross-entropy loss function is used for the classification loss L_{loc}, according to Equation (9).

$$L_{cls} = -\sum_{i=1}^{s^2}\sum_{j=1}^{B}\sum_{c\in classes} I_{i,j}^{obj}\left[\hat{p}_i^j \log\left(p_i^j\right) + \left(1 - \hat{p}_i^j\right)\log\left(1 - p_i^j\right)\right] \tag{9}$$

where S, B and $I_{i,j}^{obj}$ have the same meaning as in the context, c is the currently identified category and classes are all the classes to be detected, p_i^j and $\hat{p}_i^j$ are the predicted and true probabilities that the target in the i-th grid, j-th anchor box belongs to class c, respectively. The confidence lossL_{conf} is computed according to Equation (10).

$$\begin{aligned} L_{conf} = &-\sum_{i=0}^{s^2}\sum_{j=0}^{B} I_{i,j}^{obj}\left[\hat{C}_i^j \log\left(C_i^j\right) + \left(1 - \hat{C}_i^j\right)\log\left(1 - C_i^j\right)\right] \\ &-\lambda_{noobj}\sum_{i=1}^{s^2}\sum_{j=0}^{B} I_{i,j}^{noobj}\left[\hat{C}_i^j \log\left(C_i^j\right) + \left(1 - \hat{C}_i^j\right)\log\left(1 - C_i^j\right)\right] \end{aligned} \tag{10}$$

where, $I_{i,j}^{noobj}$ denotes the i-th grid, whether the j-th anchor box does not have a target, no target is 1, otherwise is 0; λ_{noobj} is a constant coefficient, taken as 0.5, to balance the effect of positive and negative samples on the loss function; C_i^j and $\hat{C}_i^j$ are the confidence levels of the prediction and truth boxes respectively.

3 Experimental results and analysis

The experiments were conducted using Python programming language and the PyTorch deep learning framework (version 1.7.0). Taking advantage of the server's configuration, which included two RTX 3090 GPUs, the training process efficiently utilized parallel processing. The Adam optimizer was employed to optimize the training process. A batch size of 8 was selected, striking a balance between computational efficiency and model convergence. To ensure comprehensive learning and convergence, the models were trained for 300 epochs. This experimental setup effectively maximized computational resources, enabling accurate and reliable model training.

3.1 Performance comparisons

In this paper, we evaluate the performance of disease detection models using the mean Average Precision (mAP), Precision, and Recall as metrics. The mAP is calculated by summing the Average Precision values for all categories and dividing it by the total number of categories, as shown in Equation (11).

$$AP = \int_0^1 p(r)dr \tag{11}$$

$$mAP = \frac{1}{n}\sum_{i=1}^{n} AP_i$$

where n represents the class number, AP_i denotes the Average Precision values for each category. This formulation enables a comprehensive and concise evaluation of the model's overall detection accuracy, capturing its performance across diverse disease categories.

Precision provides valuable insights into the model's capability to accurately identify and classify target frames. It quantifies the ratio of correctly identified frames to the total predicted frames, providing a measure of the model's precision and accuracy in target detection. Equation (12) represents the mathematical expression of Precision.

$$P = \frac{TP}{TP + FP} \tag{12}$$

Recall is defined as the ratio of correctly detected target frames to the total number of target frames in the dataset, assessing the model's ability to identify all instances of the target without missing any. Equation (13) represents the mathematical expression for Recall.

$$R = \frac{TP}{TP + FN} \tag{13}$$

This study evaluates the performance of TeaDiseaseNet by comparing and analyzing its detection and identification results with various classical CNN models, including SSD (Liu et al., 2016), Faster R-CNN (Ren et al., 2015) YOLOv3 (Redmon and Farhadi, 2018), YOLOv4s (Bochkovskiy et al., 2020), YOLOv5s (Jocher et al., 2022), YOLO-Tea (Xue et al., 2023), and AX-RetinaNet (Bao et al., 2022). Table 1 presents these networks' detection and recognition experiments' precision, recall, and mean Average Precision (mAP) values. The results demonstrates the outstanding detection accuracy of TeaDiseaseNet. TeaDiseaseNet achieves superior detection accuracy compared to models that employ model scaling, such as

YOLOv4s and YOLOv5s. This remarkable performance can be attributed to the utilization of DenseNet, which incorporates dense connectivity in the network, enhancing feature reuse and gradient flow. Moreover, TeaDiseaseNet employs effective techniques for multi-scale feature extraction and fusion.

It is worth noting that the YOLOv3 algorithm exhibits higher detection accuracy than YOLOv4s and YOLOv5s, potentially because of its shallower depth and smaller feature map width. The detection accuracy of Faster R-CNN is higher than that of YOLOv3 by 1.9%. This performance difference arises because Faster R-CNN is a two-stage target detection algorithm. It generates candidate regions using a region proposal network and selects the best candidate regions using a region classification network. In contrast, YOLOv3 is a one-stage target detection algorithm that predicts object locations and classes across the entire image by taking the entire image as input. Despite requiring more computational resources, Faster R-CNN delivers higher accuracy and fewer false positives compared to YOLOv3. Furthermore, TeaDiseaseNet demonstrates a slightly higher average accuracy compared to the SSD algorithm.

In conclusion, this paper presents a significant advancement in disease detection by employing CNN characteristics and incorporating a self-attentiveness mechanism. TeaDiseaseNet utilizes CNN to extract multi-scale feature maps that encompass abundant spatial information at various levels of detail. Inspired by human visual mechanisms, this design enhances the model's capability to effectively handle complex backgrounds and scale variations in disease images. The incorporation of the attention mechanism empowers TeaDiseaseNet to automatically select and prioritize the most relevant features within an image, significantly enhancing disease detection accuracy.

Figure 6 illustrates the average loss value curve of TeaDiseaseNet during training iterations. The plot demonstrates that the loss value stabilizes around 0.39 after approximately 255 iterations. The slight fluctuations observed in the loss value after convergence can be attributed to the inherent complexity and variability of the training data. The results indicate that TeaDiseaseNet has successfully learned and adapted to the training data, as evidenced by the convergence of the parameters and satisfactory performance.

3.2 Evaluation of TeaDiseaseNet

The effectiveness of the proposed TeaDiseaseNet algorithm was evaluated using the provided dataset. Table 2 presents the accuracy of the model in recognizing each tea disease. The results highlight the significant advantages of the algorithm for tea disease detection. The algorithm achieved high accuracy rates of 92.1% and recall rates of 92.9% for tea round red star disease, with an average accuracy rate of 94.5%. These findings indicate accurate identification and significant improvements in detecting this

TABLE 1 The Comparison of different networks.

Network	Year	Precision (%)	Recall (%)	mAP (%)
SSD	2016	86.5	89.1	88.4
Faster R-CNN	2015	91.5	87.3	92.2
YOLOv3	2018	94.2	84.6	90.3
YOLOv4s	2020	90.7	85.9	88.7
YOLOv5s	2020	92.3	86.5	89.4
AX-RetinaNet	2022	96.8	94	93.8
YOLO-Tea	2023	–	–	79.3
TeaDiseaseNet	2023	95.3	97.1	93.5

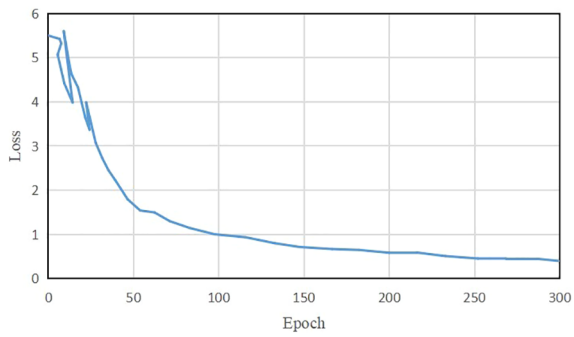


FIGURE 6
The Loss changing graph of TeaDiseaseNet.

particular disease. Tea webcake disease exhibits slightly lower accuracy and recall rates of 89.2% and 85.8%, respectively. However, the algorithm achieves an average accuracy rate of 95.4%, surpassing the individual accuracy and recall values. This demonstrates the algorithm's capability to overcome challenges related to small spot areas. The algorithm also performs remarkably well in detecting other tea diseases, including tea algae spots, tea cake disease, tea cloudy leaf blight, and tea red rust algae disease. These diseases exhibit high values across all evaluation metrics. The results demonstrate significant innovations and benefits in multi-scale tea disease detection. The algorithm achieves high recognition accuracy for large-scale tea redscab and small-scale tea exobasidium blight.

The algorithm achieves high recognition accuracy for large-scale tea redscab and small-scale tea exobasidium blight. The performance evaluation of each network was conducted based on metrics such as accuracy, recall, and average accuracy, and the results are presented in Table 3. The results revealed that the DenseNet model, serving as the backbone network, performed the best in accuracy, recall, and average accuracy. The algorithm achieves high recognition accuracy for both large-scale tea redscab and small-scale tea exobasidium blight. DenseNet exhibits strong resistance to overfitting, making it particularly suitable for scenarios with limited training data. A notable characteristic of DenseNet is its utilization of feature reuse through feature concatenation across channels. This enables DenseNet to achieve superior performance compared to ResNet-101, while utilizing fewer parameters and incurring lower computational cost. In contrast, Darknet53 is a lightweight convolutional neural network, however, it proves to be challenging to train. DenseNet performs admirably in the tea disease detection task. As a result, this paper selects DenseNet as the underlying network structure for the proposed algorithm. The algorithm effectively resolves the scaling issue by establishing a multi-scale feature representation and enhances overall performance. In summary, the algorithm proposed in this study demonstrates improved accuracy compared to other models, thereby representing significant progress in the field of tea disease detection.

3.3 Ablation studies

To validate the effectiveness of the proposed network model, incremental ablation experiments were conducted. Each network module was incrementally incorporated based on the DenseNet backbone architecture. This approach allowed for a comprehensive evaluation of each module's contribution to the overall performance. This step-by-step approach aimed to enable a comprehensive evaluation of the individual contribution of each module to the overall performance.

TABLE 2 Performance in detecting different tea diseases.

Tea Disease	P (%)	R (%)	mAP (%)
Tea exobasidium blight	89.2	85.8	95.4
Tea red scab	92.1	92.9	94.5
Tea algae leaf spot	94.9	88.8	93.5
Tea cake	90.0	91.4	94.7
Tea cloud leaf blight	88.5	89.5	92.0
Tea red rust	85.4	87.7	90.9

TABLE 3 Performance comparison of different backbone networks.

Backbone	P (%)	R (%)	mAP (%)
DenseNet	95.3	97.1	93.5
Resnet-101	91.3	90.2	92.8
Darknet53	91.8	90.5	93.2

The results of the ablation experiments conducted for each module are presented in Table 4. Including of the MFEM+SSAM module results in a substantial performance improvement, with a 2.2% increase in mAP compared to using the MFEM module alone. This improvement can be attributed to utilizing the multi-head self-attention mechanism within the MFEM+SSAM module. This mechanism captures global contextual information from multi-scale feature maps and facilitates the interaction between global and local information. Assigning weights to features, such as spot color and leaf edge, enhances the detection accuracy. Furthermore, the new scale-aware feature fusion (SFF) module adopts a channel attention mechanism to fuse features of different scales. It focuses on the feature channels containing discriminative information and assigns a higher weight distribution to them, effectively improving the detection performance (Chen et al., 2020). The SFF module effectively fuses information from tea disease features of various scales, resulting in improved accuracy of localization and identification. The introduction of the SFF module enhances the mAP by 0.7%, indicating its contribution to improved detection accuracy of the network.

3.4 Visualization and discussion

Representative disease images were selected to showcase the exceptional performance of the proposed model in effectively addressing challenges posed by continuous scale variations and complex backgrounds. This visualization demonstrates the model's ability to detect diseases accurately. Figures 7 and 8 present the original disease images on the left and the model's detection results on the right. Rectangular boxes indicate the identified disease types and their corresponding confidence levels.

Figure 7 demonstrates the model's ability to effectively identify diseases with varying scales and address disease scenarios characterized by continuous scale variations. This showcases the

TABLE 4 Results of ablation experiments.

Backbone (MFEM)	SSAM	SFF	mAP (%)
√			90.6
√	√		92.8
√	√	√	93.5

FIGURE 7
Disease scenarios with continuous scale changes.

algorithm's capacity to extract rich global contextual information at multiple scales and accurately detect scale variations by comparing global and local information. Figure 8 highlights the model's effectiveness in eliminating complex background interferences, such as branches and fallen leaves in disease scenes. This can be attributed to the feature extraction and fusion networks, which enable the proposed method to accurately detect disease areas within complex scenes by capturing dependencies between input feature scales.

Additionally, a comparison was conducted between the YOLOv3 and TeaDiseaseNet models using images of tea leaf diseases, as depicted in Figure 9. The YOLOv3 model exhibited missed detections and inaccurate annotation box positions, whereas TeaDiseaseNet accurately detected and confidently annotated the diseases. The superior performance of TeaDiseaseNet can be attributed to its multi-scale self-attention mechanism, which enhances the acquisition of semantic and location information in the images. This results in improved feature extraction and detection accuracy.

The results demonstrate that TeaDiseaseNet outperforms YOLOv3 in terms of detection accuracy and robustness, owing to its enhanced feature extraction capabilities and multi-scale self-attention mechanism.

4 Conclusion

This paper introduces TeaDiseaseNet, a novel tea disease detection model that effectively addresses challenges posed by complex backgrounds and variable scales. By incorporating a multi-scale self-attentive mechanism, TeaDiseaseNet enables effective interactions between global and local features across multiple scales. This mitigates the impact of scale variations and complex backgrounds on detection accuracy. Experimental results demonstrate that TeaDiseaseNet surpasses state-of-the-art methods, exhibiting exceptional accuracy and robustness in scale variations and background interference scenarios. These findings provide valuable insights for intelligent tea disease diagnosis, supporting tea farmers with accurate detection capabilities and enabling timely control measures to protect tea plantations, improve tea quality, and enhance yields.

In addition to the benefits and contributions highlighted in the conclusion, this study also has certain limitations that need to be acknowledged. Firstly, the use of a limited dataset may not fully capture the diversity of tea diseases. Including a wider range of tea diseases would enhance the representativeness and applicability of

FIGURE 8
Disease scenes with complex backgrounds.

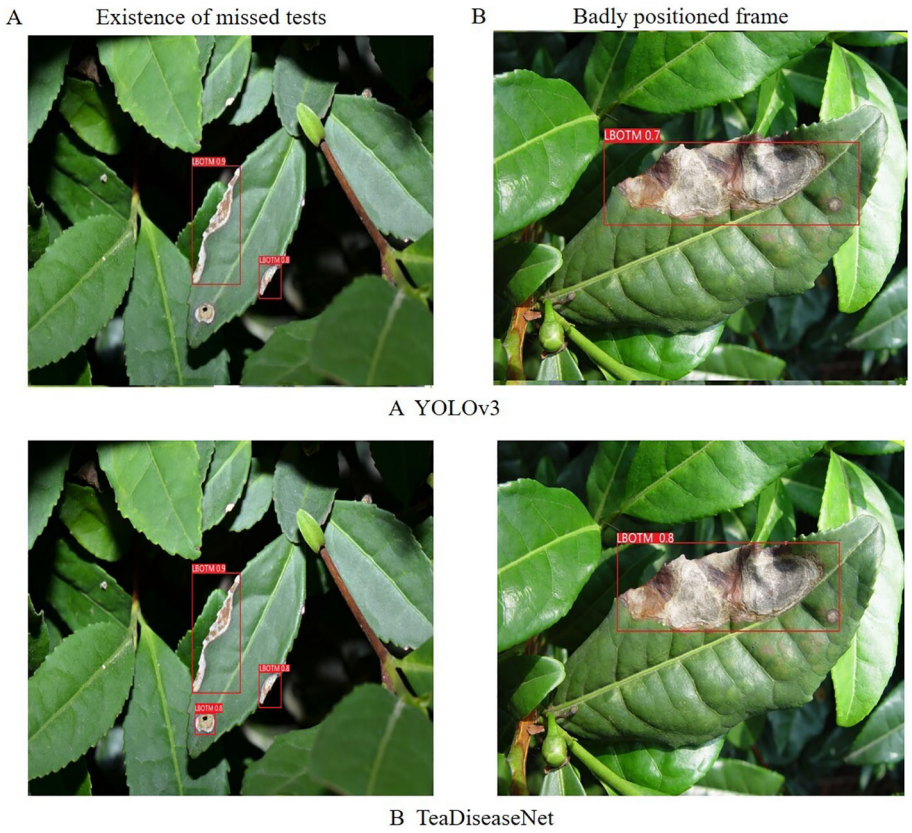


FIGURE 9
The results of algorithm **(A)** YOLOv3 and algorithm **(B)** TeaDiseaNet, in the existence of missed tests and badly positioned frame scenarios, respectively.

the detection system. Secondly, biases in the training data, such as imbalances in disease instances or variations introduced by different image acquisition systems, could affect the performance of the tea disease detection system. Efforts should be made to address these biases and enhance the system's robustness. Additionally, the study focuses on offline detection, which may not be practical for real-time implementation in tea plantations. Future research should explore real-time implementation, taking into account the resource and time constraints associated with practical deployment. Lastly, interpreting the decision-making processes of the deep learning model is challenging due to their complexity. Enhancing the interpretability of the model would enhance its usefulness in decision-making for tea farmers. Addressing these limitations can improve the practicality and effectiveness of tea disease detection systems.

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. The raw data can be accessed at the following link: https://www.jianguoyun.com/p/DRwyMxYQqJnmCxiGl5IFIAA.

Author contributions

YS: Writing – original draft, Supervision. FW: Methodology, Data curation, Writing – review & editing. HG: Investigation, Funding acquisition, Writing – review & editing. RL: Writing – review & editing, Conceptualization. JY: Writing – review & editing, Resources, Visualization. JS: Writing – review & editing, Software.

Funding

The author(s) declare financial support was received for the research, authorship, and/or publication of this article. This work is supported by National Natural Science Foundation of China (62062004), Natural Science Foundation of Henan Province (222300420274), Academic Degrees & Graduate Education Reform Project of Henan Province (2021SJGLX057Y), Postgraduate Education Reform and Quality Improvement Project of Henan Province (YJS2023SZ23), Nanhu Scholars Program for Young Scholars of XYNU.

Conflict of interest

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

Publisher's note

References

Abade, A., Ferreira, P. A., and de Barros Vidal, F. (2021). Plant diseases recognition on images using convolutional neural networks: A systematic review. *Comput. Electron. Agric.* 185, 106125. doi: 10.1016/j.compag.2021.106125

Akanksha, E., Sharma, N., and Gulati, K. (2021). "OPNN: optimized probabilistic neural network based automatic detection of maize plant disease detection," in *2021 6th international conference on inventive computation technologies (ICICT)*. (Piscataway, NJ:IEEE) 1322–1328.

Alirezazadeh, P., Schirrmann, M., and Stolzenburg, F. (2023). Improving deep learning-based plant disease classification with attention mechanism. *Gesunde Pflanzen* 75 (1), 49–59. doi: 10.1007/s10343-022-00796-y

Alruwaili, M., Siddiqi, M. H., Khan, A., Azad, M., Khan, A., and Alanazi, S. (2022). RTF-RCNN: An architecture for real-time tomato plant leaf diseases detection in video streaming using Faster-RCNN. *Bioengineering* 9 (10), 565. doi: 10.3390/bioengineering9100565

Ashwinkumar, S., Rajagopal, S., Manimaran, V., and Jegajothi, B. (2022). Automated plant leaf disease detection and classification using optimal MobileNet based convolutional neural networks. *Mater. Today: Proc.* 51, 480–487. doi: 10.1016/j.matpr.2021.05.584

Ayan, E., Erbay, H., and Varçın, F. (2020). Crop pest classification with a genetic algorithm-based weighted ensemble of deep convolutional neural networks. *Comput. Electron. Agric.* 179, 105809. doi: 10.1016/j.compag.2020.105809

Bao, W., Fan, T., Hu, G., Liang, D., and Li, H. (2022). Detection and identification of tea leaf diseases based on AX-RetinaNet. *Sci. Rep.* 12 (1), 1–16. doi: 10.1038/s41598-022-06181-z

Bhavsar, N., Tayade, A., Ghodmare, H., Wagdharikar, N., De, S., and Raj, A. B. (2022) A review on Image Processing Algorithm for Foliage Target Detection and classification. *Research Publish Journals*. 10, 52–61. doi: 10.5281/zenodo.7486512

Bochkovskiy, A., Wang, C.-Y., and Liao, H.-Y. M. (2020). Yolov4: Optimal speed and accuracy of object detection. *arXiv* 2004, 10934. doi: 10.48550/arXiv.2004.10934

Borhani, Y., Khoramdel, J., and Najafi, E. (2022). A deep learning based approach for automated plant disease classification using vision transformer. *Sci. Rep.* 12 (1), 11554. doi: 10.1038/s41598-022-15163-0

Carion, N., Massa, F., Synnaeve, G., Usunier, N., Kirillov, A., and Zagoruyko, S. (2020). "End-to-end object detection with transformers," in *European conference on computer vision*, Glasgow, UK, August 23–28, 2020. (Berlin, German: Springer) 213–229.

Chen, J., Liu, Q., and Gao, L. (2019). Visual tea leaf disease recognition using a convolutional neural network model. *Symmetry* 11 (3), 343. doi: 10.3390/sym11030343

Chen, X., Zhou, G., Chen, A., Yi, J., Zhang, W., and Hu, Y. (2020). Identification of tomato leaf diseases based on combination of ABCK-BWTR and B-ARNet. *Comput. Electron. Agric.* 178, 105730. doi: 10.1016/j.compag.2020.105730

Debnath, B., Haldar, D., and Purkait, M. K. (2021). Potential and sustainable utilization of tea waste: A review on present status and future trends. *J. Environ. Chem. Eng.* 9 (5), 106179. doi: 10.1016/j.jece.2021.106179

Dhaka, V. S., Meena, S. V., Rani, G., Sinwar, D., Ijaz, M. F., and Woźniak, M. (2021). A survey of deep convolutional neural networks applied for prediction of plant leaf diseases. *Sensors* 21 (14), 4749. doi: 10.3390/s21144749

Drew, L. (2019). The growth of tea. *Nature* 566 (7742), S2–S2. doi: 10.1038/d41586-019-00395-4

Guo, M.-H., Xu, T.-X., Liu, J.-J., Liu, Z.-N., Jiang, P.-T., Mu, T.-J., et al. (2022). Attention mechanisms in computer vision: A survey. *Comput. Visual Media* 8 (3), 331–368. doi: 10.1007/s41095-022-0271-y

Hou, J., Yang, C., He, Y., and Hou, B. (2023). Detecting diseases in apple tree leaves using FPN–ISResNet–Faster RCNN. *Eur. J. Remote Sens.* 56 (1), 2186955. doi: 10.1080/22797254.2023.2186955

Hu, Y., Deng, X., Lan, Y., Chen, X., Long, Y., and Liu, C. (2023). Detection of rice pests based on self-attention mechanism and multi-scale feature fusion. *Insects* 14 (3), 280. doi: 10.3390/insects14030280

Hu, G., Wang, H., Zhang, Y., and Wan, M. (2021a). Detection and severity analysis of tea leaf blight based on deep learning. *Comput. Electrical Eng.* 90, 107023. doi: 10.1016/j.compeleceng.2021.107023

Hu, G., Wei, K., Zhang, Y., Bao, W., and Liang, D. (2021b). Estimation of tea leaf blight severity in natural scene images. *Precis. Agric.* 22, 1239–1262. doi: 10.1007/s11119-020-09782-8

Jiang, F., Lu, Y., Chen, Y., Cai, D., and Li, G. (2020). Image recognition of four rice leaf diseases based on deep learning and support vector machine. *Comput. Electron. Agric.* 179, 105824. doi: 10.1016/j.compag.2020.105824

Jiao, L., Zhang, R., Liu, F., Yang, S., Hou, B., Li, L., et al. (2021). New generation deep learning for video object detection: A survey. *IEEE Trans. Neural Networks Learn. Syst.* 33 (8), 3195–3215. doi: 10.1109/TNNLS.2021.3053249

Jocher, G., Chaurasia, A., Stoken, A., Borovec, J., Kwon, Y., Michael, K., et al. (2022). ultralytics/yolov5: V7. 0-YOLOv5 SOTA realtime instance segmentation. *Zenodo*. doi: 10.5281/zenodo.7347926

Khan, S., Naseer, M., Hayat, M., Zamir, S. W., Khan, F. S., and Shah, M. (2022). Transformers in vision: A survey. ACM computing surveys (CSUR) 54 (10s) 1-41.

Kirti, and Rajpal, N. (2023). A multi-crop disease identification approach based on residual attention learning. *J. Intelligent Syst.* 32 (1), 20220248. doi: 10.1515/jisys-2022-0248

Kirti, Rajpal, N., and Vishwakarma., P. V. (2023). Automatic adaptive weighted fusion of features-based approach for plant disease identification. *J. Intelligent Syst.* 32 (1), 20220247. doi: 10.1515/jisys-2022-0247

Krisnandi, D., Pardede, H. F., Yuwana, R. S., Zilvan, V., Heryana, A., Fauziah, F., et al. (2019). Diseases classification for tea plant using concatenated convolution neural network. *CommIT (Commun. Inf. Technol.) J.* 13 (2), 67–77. doi: 10.21512/commit.v13i2.5886

Lanjewar, M. G., and Panchbhai, K. G. (2023). Convolutional neural network based tea leaf disease prediction system on smart phone using paas cloud. *Neural Comput. Appl.* 35 (3), 2755–2771. doi: 10.1007/s00521-022-07743-y

Latha, R., Sreekanth, G., Suganthe, R., Rajadevi, R., Karthikeyan, S., Kanivel, S., et al. (2021). "Automatic detection of tea leaf diseases using deep convolution neural network," in *2021 International Conference on Computer Communication and Informatics (ICCCI)*. (Piscataway, NJ:IEEE) 1–6.

Lee, S. H., Lin, S. R., and Chen, S. F. (2020). Identification of tea foliar diseases and pest damage under practical field conditions using a convolutional neural network. *Plant Pathol.* 69 (9), 1731–1739. doi: 10.1111/ppa.13251

Li, L., Zhang, S., and Wang, B. (2021). Plant disease detection and classification by deep learning—a review. *IEEE Access* 9, 56683–56698. doi: 10.1109/ACCESS.2021.3069646

Lin, J., Bai, D., Xu, R., and Lin, H. (2023). TSBA-YOLO: An improved tea diseases detection model based on attention mechanisms and feature fusion. *Forests* 14 (3), 619. doi: 10.3390/f14030619

Liu, W., Anguelov, D., Erhan, D., Szegedy, C., Reed, S., Fu, C.-Y., et al. (2016). "Ssd: Single shot multibox detector," in *Computer Vision–ECCV 2016: 14th European Conference*, Amsterdam, The Netherlands, October 11–14, 2016. (Berlin, German: Springer) 21–37.

Liu, B.-Y., Fan, K.-J., Su, W.-H., and Peng, Y. (2022). Two-stage convolutional neural networks for diagnosing the severity of alternaria leaf blotch disease of the apple tree. *Remote Sens.* 14 (11), 2519. doi: 10.3390/rs14112519

Liu, L., Wang, R., Xie, C., Yang, P., Wang, F., Sudirman, S., et al. (2019). PestNet: An end-to-end deep learning approach for large-scale multi-class pest detection and classification. *IEEE Access* 7, 45301–45312. doi: 10.1109/ACCESS.2019.2909522

Lu, J., Tan, L., and Jiang, H. (2021). Review on convolutional neural network (CNN) applied to plant leaf disease classification. *Agriculture* 11 (8), 707. doi: 10.3390/agriculture11080707

Mahum, R., Munir, H., Mughal, Z.-U.-N., Awais, M., Sher Khan, F., Saqlain, M., et al. (2023). A novel framework for potato leaf disease detection using an efficient deep learning model. *Hum. Ecol. Risk Assessment: Int. J.* 29 (2), 303–326. doi: 10.1080/10807039.2022.2064814

Mukhopadhyay, S., Paul, M., Pal, R., and De, D. (2021). Tea leaf disease detection using multi-objective image segmentation. *Multimedia Tools Appl.* 80, 753–771. doi: 10.1007/s11042-020-09567-1

Pandian, J. A., Nisha, S. N., Kanchanadevi, K., Pandey, A. K., and Rima, S. K. (2023). Grey Blight Disease Detection on Tea Leaves Using Improved Deep Convolutional Neural Network. *Computational Intelligence and Neuroscience* 2023. doi: 10.1155/2023/7876302

Prabu, S., TapasBapu, B., Sridhar, S., and Nagaraju, V. (2022). "Tea plant leaf disease identification using hybrid filter and support vector machine classifier technique," in *Advances in Data Science and Management: Proceedings of ICDSM 2021* (Berlin, German:Springer), 591–601.

Qi, C., Gao, J., Pearson, S., Harman, H., Chen, K., and Shu, L. (2022). Tea chrysanthemum detection under unstructured environments using the TC-YOLO model. *Expert Syst. Appl.* 193, 116473. doi: 10.1016/j.eswa.2021.116473

Redmon, J., and Farhadi, A. (2018). Yolov3: An incremental improvement. *arXiv* 1804, 02767. doi: 10.48550/arXiv.1804.02767

Ren, S., He, K., Girshick, R., and Sun, J. (2015). Faster r-cnn: Towards real-time object detection with region proposal networks. *Adv. Neural Inf. Process. Syst.* 28. doi: 10.1109/TPAMI.2016.2577031

Roy, A. M., and Bhaduri, J. (2022). Real-time growth stage detection model for high degree of occultation using DenseNet-fused YOLOv4. Computers and Electronics in Agriculture 193 106694.

Sapna, N., Rajni, J., Sudeep, M., Alka, A., Ashraful, H. M., Akshay, D., et al. (2023). Deep transfer learning model for disease identification in wheat crop. *Ecol. Inf.* 75, 102068. doi: 10.1016/j.ecoinf.2023.102068

Steven, S. (2021). Tea leaf pest detection using Support Vector Machine (SVM) METHOD IN PTPN IV unit bah butong. *INFOKUM* 9 (2, June), 299–305. https://seaninstitute.org/infor/index.php/infokum/article/view/127

Sudhesh, K. M., Sowmya, V., Kurian, P. S., and Sikha, O. K. (2023). AI based rice leaf disease identification enhanced by Dynamic Mode Decomposition. *Eng. Appl. Artif. Intell.* 120, 105836. doi: 10.1016/j.engappai.2023.105836

Sunil, C., Jaidhar, C., and Patil, N. (2023). Tomato plant disease classification using multilevel feature fusion with adaptive channel spatial and pixel attention mechanism. *Expert Syst. Appl.* 228, 120381. doi: 10.1016/j.eswa.2023.120381

Tetila, E. C., MaChado, B. B., Astolfi, G., de Souza Belete, N. A., Amorim, W. P., Roel, A. R., et al. (2020). Detection and classification of soybean pests using deep learning with UAV images. *Comput. Electron. Agric.* 179, 105836. doi: 10.1016/j.compag.2020.105836

Tholkapiyan, M., Aruna Devi, B., Bhatt, D., Saravana Kumar, E., Kirubakaran, S., and Kumar, R. (2023). Performance analysis of rice plant diseases identification and classification methodology. *Wireless Pers. Commun.* 130 (2), 1317–1341. doi: 10.1007/s11277-023-10333-3

Wang, D., Wang, J., Li, W., and Guan, P. (2021). T-CNN: Trilinear convolutional neural networks model for visual detection of plant diseases. *Comput. Electron. Agric.* 190, 106468. doi: 10.1016/j.compag.2021.106468

Wang, F., Wang, R., Xie, C., Yang, P., and Liu, L. (2020). Fusing multi-scale context-aware information representation for automatic in-field pest detection and recognition. *Comput. Electron. Agric.* 169, 105222. doi: 10.1016/j.compag.2020.105222

Xiong, Y., Liang, L., Wang, L., She, J., and Wu, M. (2020). Identification of cash crop diseases using automatic image segmentation algorithm and deep learning with expanded dataset. *Comput. Electron. Agric.* 177, 105712. doi: 10.1016/j.compag.2020.105712

Xu, L., Cao, B., Ning, S., Zhang, W., and Zhao, F. (2023). Peanut leaf disease identification with deep learning algorithms. *Mol. Breed.* 43 (4), 25. doi: 10.1007/s11032-023-01370-8

Xue, Z., Xu, R., Bai, D., and Lin, H. (2023). YOLO-tea: A tea disease detection model improved by YOLOv5. *Forests* 14 (2), 415. doi: 10.3390/f14020415

Yang, L., Yu, X., Zhang, S., Long, H., Zhang, H., Xu, S., et al. (2023). GoogLeNet based on residual network and attention mechanism identification of rice leaf diseases. *Comput. Electron. Agric.* 204, 107543. doi: 10.1016/j.compag.2022.107543

Yogeshwari, M., and Thailambal, G. (2021). Automatic feature extraction and detection of plant leaf disease using GLCM features and convolutional neural networks. *Mater. Today: Proc* 81, 530–536. doi: 10.1016/j.matpr.2021.03.700

Yu, H., Liu, J., Chen, C., Heidari, A. A., Zhang, Q., Chen, H., et al. (2021). Corn leaf diseases diagnosis based on K-means clustering and deep learning. *IEEE Access* 9, 143824–143835. doi: 10.1109/ACCESS.2021.3120379

Zha, M., Qian, W., Yi, W., and Hua, J. (2021). A lightweight YOLOv4-Based forestry pest detection method using coordinate attention and feature fusion. *Entropy* 23 (12), 1587. doi: 10.3390/e23121587

Zheng, S., Lu, J., Zhao, H., Zhu, X., Luo, Z., Wang, Y., et al. (2021). "Rethinking semantic segmentation from a sequence-to-sequence perspective with transformers," in *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*. (Piscataway, NJ: IEEE) 6881–6890.

Zhou, H., Fang, Z., Wang, Y., and Tong, M. (2023). Image generation of tomato leaf disease identification based on small-ACGAN. *Computers Mater. Continua* 76 (1), 175–194. doi: 10.32604/cmc.2023.037342

Zhu, W., Sun, J., Wang, S., Shen, J., Yang, K., and Zhou, X. (2022). Identifying field crop diseases using transformer-embedded convolutional neural network. *Agriculture* 12 (8), 1083. doi: 10.3390/agriculture12081083

Zou, X., Ren, Q., Cao, H., Qian, Y., and Zhang, S. (2020). Identification of tea diseases based on spectral reflectance and machine learning. *J. Inf. Process. Syst.* 16 (2), 435–446. doi: 10.3745/jips.02.0133

frontiers | Frontiers in Plant Science

TYPE Original Research
PUBLISHED 12 October 2023
DOI 10.3389/fpls.2023.1275004

OPEN ACCESS

EDITED BY
Xi Tian,
Beijing Academy of Agriculture and Forestry Sciences, China

REVIEWED BY
Ting An,
Shandong Academy of Agricultural Sciences, China
Kusumiyati Kusumiyati,
Padjadjaran University, Indonesia
Yi Shilai,
Southwest University, China

*CORRESPONDENCE
Guohui Han
hghui2007@126.com

RECEIVED 09 August 2023
ACCEPTED 18 September 2023
PUBLISHED 12 October 2023

CITATION
Li X, Peng F, Wei Z, Han G and Liu J (2023) Non-destructive detection of protein content in mulberry leaves by using hyperspectral imaging.
Front. Plant Sci. 14:1275004.
doi: 10.3389/fpls.2023.1275004

Non-destructive detection of protein content in mulberry leaves by using hyperspectral imaging

Xunlan Li, Fangfang Peng, Zhaoxin Wei, Guohui Han* and Jianfei Liu

Research Institute of Pomology, Chongqing Academy of Agricultural Sciences, Chongqing, China

Protein content is one of the most important indicators for assessing the quality of mulberry leaves. This work is carried out for the rapid and non-destructive detection of protein content of mulberry leaves using hyperspectral imaging (HSI) (Specim FX10 and FX17, Spectral Imaging Ltd., Oulu, Finland). The spectral range of the HSI acquisition system and data processing methods (pretreatment, feature extraction, and modeling) is compared. Hyperspectral images of three spectral ranges in 400–1,000 nm (Spectral Range I), 900–1,700 nm (Spectral Range II), and 400–1,700 nm (Spectral Range III) were considered. With standard normal variate (SNV), Savitzky–Golay first-order derivation, and multiplicative scatter correction used to preprocess the spectral data, and successive projections algorithm (SPA), competitive adaptive reweighted sampling, and random frog used to extract the characteristic wavelengths, regression models are constructed by using partial least square and least squares-support vector machine (LS-SVM). The protein content distribution of mulberry leaves is visualized based on the best model. The results show that the best results are obtained with the application of the model constructed by combining SNV with SPA and LS-SVM, showing an R^2 of up to 0.93, an RMSE of just 0.71 g/100 g, and an RPD of up to 3.83 based on the HSI acquisition system of 900–1700 nm. The protein content distribution map of mulberry leaves shows that the protein of healthy mulberry leaves distributes evenly among the mesophyll, with less protein content in the vein of the leaves. The above results show that rapid, non-destructive, and high-precision detection of protein content of mulberry leaves can be achieved by applying the SWIR HSI acquisition system combined with the SNV-SPA-LS-SVM algorithm.

KEYWORDS

hyperspectral imaging, mulberry leaf, protein content, non-destructive detection, visible and near-infrared

1 Introduction

Mulberry leaves are rich in a variety of bioactive ingredients necessary for the human body, with such functions as anti-obesity (Li et al., 2019), anti-oxidation and antibacterial (Thabti et al., 2014), and anti-diabetes (Riche et al., 2017), and thus are considered to be beneficial in the Asian population. Compared to animal protein, the abundant protein of mulberry leaves contains no animal cholesterol, with an amino acid pattern similar to that of the human body (Gryn-Rynko et al., 2016; Sun et al., 2017). In recent years, mulberry leaves are eaten as a vegetable, and used as a traditional source of animal feed protein as well in Asian countries (Srivastava et al., 2003; Yu et al., 2018). The protein content is one of the most important indicators for assessing the quality of mulberry leaves used as an animal feed source or a fresh vegetable.

At present, the methods for determining protein content in leaves are mainly chemical analysis methods (Ledoux and Lamy, 1986; Chromý et al., 2015; Denholm et al., 2021), such as the Kjeldahl nitrogen determination method. Such methods require the samples to undergo not only drying, grinding, and other destructive treatments, but also deboiling, distillation, and titration under the condition of concentrated sulfuric acid being added. This is a complex process producing chemical pollution. In light of this, it is highly necessary to introduce a non-destructive and rapid determination of protein content of mulberry leaves.

Hyperspectral imaging (HSI) combining imaging technology with spectral technology can provide both spectral and spatial information of substances. With the advantages of non-destructiveness, high efficiency, and low cost, HSI is widely used in non-destructive detection of protein content of different farm products, including meat products such as pork (R^2_P = 0.9161 and RMSEP = 2.71 mg/g) (Ma J. et al., 2019), lamb (R^2_p = 0.67 and RMSEP = 0.41) (Pu et al., 2014), and beef (R^2_P = 0.86 and SEP = 0.29) (ElMasry et al., 2013), and grain products such as wheat (R^2_P = 0.79 and RMSEP = 0.94) (Caporaso et al., 2018), rice (R^2_P = 0.8011 and RMSEP = 0.52) (Ma et al., 2021), and peanuts (R^2_P = 0.912 and RMSEP = 0.438) (Cheng et al., 2017). There are studies showing that N-H bonds in proteins present absorption peaks at 1,460–1,570 nm and 2,000–2,180 nm (Shenk et al., 2007; Chelladurai and Jayas, 2014), which lead to the non-destructive detection of proteins to be conducted by mainly using the Short-Wave Infrared (SWIR) HSI system with an acquisition wavelength range of 1,000–1,700 nm or 900–2,500 nm. There are also some other researchers using visible near-infrared (Vis-NIR) HSI with an acquisition wavelength range of 400–1,000 nm for non-destructive detection of proteins of meat (Ma J. et al., 2019), rice (Onoyama et al., 2018), milk (Jin et al., 2022), and rape leaves (Zhang et al., 2015), with good results obtained. As the main parts of optical imaging systems, detectors are meant for detecting and measuring the radiation reflected or transmitted by objects. A detector made of a certain material can only detect certain wavelength ranges, and the prices of detectors vary greatly. Currently, silicon detectors (300–1,100 nm) are the most widely used Vis-NIR detectors, and their prices are very low, compared with the slightly more expensive InGaAs detectors (900–2,500 nm) and the much more expensive HgCdTe detectors (1,000–2,600 nm). At present, there are only a few studies on the non-destructive detection of proteins of mulberry leaves. Ma et al. used a 900-1,600 nm handheld near-infrared spectrometer to detect proteins of dry mulberry leaves, and by combining with partial least squares (PLS) regression and the wavelength optimization method, they obtained a prediction set R^2 of up to 0.92 (Ma Y. et al., 2019). However, this method requires the mulberry leaves to undergo drying and grinding, and the obtained data are single point data, thus leading to failure to obtain the protein content of the whole leaves. The vibrational characteristics of different molecules and functional groups vary, resulting in differences in sensitivity to specific wavelengths among different substances. Therefore, we are not clear about the best detector material and the spectral range for conducting the non-destructive detection of protein content of mulberry leaves. As a result, it is necessary to choose an optimal compromise.

This study aims at developing a non-destructive and rapid method for the detection of protein content of mulberry leaves. The main research contents are as follows: (1) analyzing the spectral characteristics of mulberry leaves at Vis-NIR (400–1,000 nm) and SWIR (900–1,700 nm); (2) comparing different pretreatment, feature extraction, and modeling methods and selecting the best optimal data processes and methods; (3) selecting the best spectral range of HSI acquisition system for the detection of protein content of mulberry leaves; and (4) visualizing the distribution of protein content of mulberry leaves by using the optimal model.

2 Materials and methods

2.1 Materials

The healthy mulberry leaves, randomly collected and washed with tap water when brought to the laboratory, undergo hyperspectral images collection and protein content determination after the surfaces of the leaves become dry. In this study, 193 samples are randomly divided into the training set and the testing set at the ratio of 7:3, with 135 and 58 samples, respectively. Among them, the training set is used for training the model, with the 10-fold cross-validation method applied to the training set to adjust the model parameters and select the optimal model, while the test set is used for assessing the final model.

2.2 Acquisition and calibration of hyperspectral images

The HSI acquisition system consists of two hyperspectral cameras (FX10 and FX17, Spectral Imaging Ltd., Oulu, Finland), the electric linear platform (Spectral Imaging Ltd., Oulu, Finland), two light sources (each light source consists of three 20-W halogen lamps), and a laptop (Figure 1). The FX10 spectral camera (Si detector) is used for acquiring hyperspectral images of the Vis-NIR region (400–1,000 nm). The FX17 spectral camera (InGaAs detector) is used for acquiring hyperspectral images of the SWIR region (900–1,700 nm).

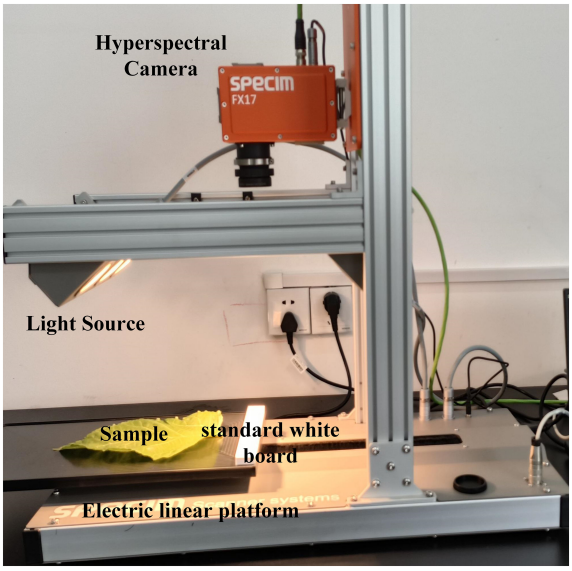


FIGURE 1
The hyperspectral imaging system.

The two light sources are at an angle of 45° of the moving platform, and the distance between the lens and the platform is 33 cm. When FX10 is used for hyperspectral image acquisition, the exposure time is set to 12.5 ms, the frame rate is 49.83 Hz, the spectral interval is 2, the spatial interval is 1, and the mobile platform moving speed is set to 11.9 mm/s. When FX17 is used for hyperspectral image acquisition, the exposure time is set to 6 ms, the frame rate is 40.5 Hz, the spectral interval is 1, the spatial interval is 1, and the moving speed of the mobile platform is set to 14.8 mm/s. The hyperspectral image acquisition is conducted after the preheating of 20 min. The white reference image W is obtained by screening the standard white board with a reflectance of 99% placed in front of the sample. The dark reference image D is obtained by screening with the lens closed. The reference images are acquired together with the hyperspectral image of the sample.

To avoid the effect caused by uneven light source intensity distribution and dark current during the image collecting process, hyperspectral image calibration is conducted (Figure 2). The following formula is used for hyperspectral image calibration.

$$R_\lambda = \frac{I_\lambda - D_\lambda}{W_\lambda - D_\lambda}$$

where R_λ is the calibrated image, I_λ is the raw image, W_λ is the white reference image, and D_λ is the dark reference image.

2.3 Determination of protein content of mulberry leaves

After the hyperspectral image acquisition, the mulberry leaves underwent drying in the oven at 105°C for 15 min and then drying at 50°C for 2 h. With the main vein removed, the leaves were ground with a mortar and passed through a 60-mesh sieve, and the protein content of mulberry leaves was determined by using Kaye nitrogen determination (Chromý et al., 2015). The sample (0.3 g) was weighed and transferred into a digestion tube. Then, 0.4 g of copper sulfate, 6 g of potassium sulfate, and 20 mL of sulfuric acid were added to the tube for digestion. Once the temperature of the digestion furnace reached 420°C, the digestion process was continued for 1 h. After the liquid in the digestion tube exhibited a green and transparent appearance, the tube was carefully removed from the furnace and allowed to cool. Once cooled, 50 mL of water was added to the tube. In the Kjeldahl nitrogen analyzer, sodium hydroxide solution, hydrochloric acid standard solution, and boric acid solution containing mixed indicators were first added. Finally, the automated Kjeldahl nitrogen analyzer was utilized to automatically perform the processes of sample addition, distillation, titration, and data recording. The protein content in the mulberry leaf can then be calculated using the provided formula.

$$X = \frac{(V1 - V2) \star C \star 0.0140 \star F \star 100}{m \star V3/100}$$

In the formula, X represents the measured protein content, $V1$ represents the volume of consumed hydrochloric acid standard solution, $V2$ represents the volume of blank consumed hydrochloric acid standard solution, $V3$ represents the volume of extracted liquid. C = 0.05 mol/L represents the concentration of hydrochloric acid standard solution. m represents the weight of the sample taken. F represents the conversion factor of nitrogen to protein, and F is taken as 6.25. 100 is the conversion factor.

2.4 Data processing

2.4.1 Region of interest identification and spectrum extraction

In this study, a whole mulberry leaf is the region of interest for spectral extraction. A gray image is obtained at 800 nm and 1,000 nm of the Vis-NIR and SWIR hyperspectral images, respectively. The Otsu method automatically calculates the segmentation threshold between the leaf and the background in the gray image, from which a binary image is obtained. Then, the ROI is obtained by conducting mask processing. Finally, the average spectral reflectance of the whole mulberry leaf at each wavelength is calculated.

2.4.2 Spectral pretreatment

In light of the high noises in the first and last bands of the original spectral data, spectral data within the ranges of 423–975 nm (Spectral Range I), 970–1,684 nm (Spectral Range II), and 423–1,684 nm (Spectral Range III) are selected for subsequent analyses. The raw spectra need to be pre-treated to eliminate the scattering caused by uneven particle distribution and different particle sizes and the influence of optical path difference on the spectral data. The standard normal variate (SNV) (Barnes et al., 1989), Savitzky–Golay combined first-order derivation (Savitzky and Golay, 1964), and multiplicative scatter correction (MSC) (Isaksson and Næs, 1988) are the commonly used spectral preprocessing methods, and

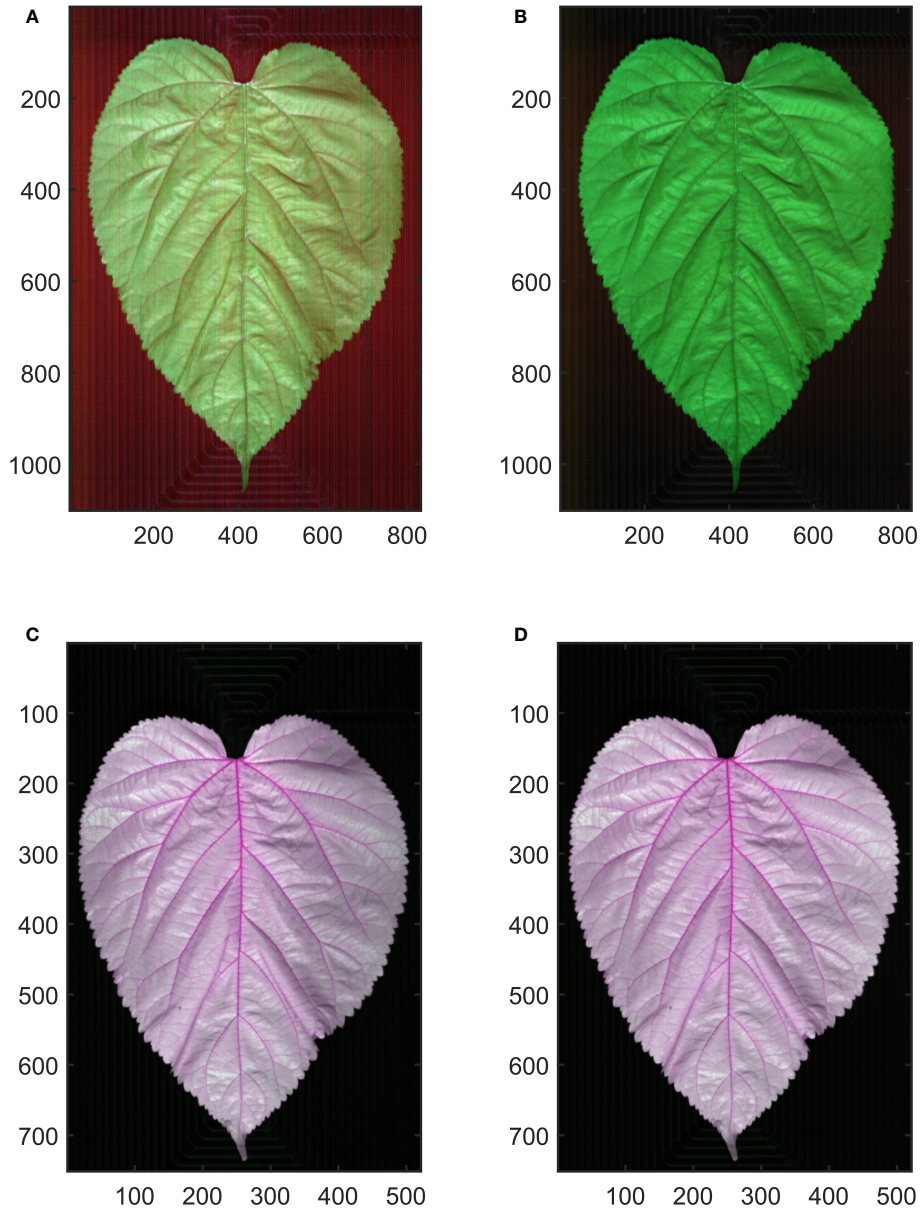


FIGURE 2
Hyperspectral image calibration. **(A, C)** show the raw hyperspectral image of Vis-NIR and SWIR region, respectively. **(B, D)** show the calibrated hyperspectral image of Vis-NIR and SWIR region, respectively.

have been shown to be effective in eliminating or reducing interference such as electrical noise, sample background, and stray light during acquisition. In order to determine the best pretreatment of spectral data, the SNV, Savitzky–Golay combined first-order derivation, and MSC are adopted in this study.

2.4.3 Variable selection

Because of the high correlation between adjacent spectral bands, successive projections algorithm (SPA), competitive adaptive reweighted sampling (CARS), and random frog (RF) are respectively used to extract characteristic wavelengths in this study to reduce model input variables and improve model efficiency and prediction accuracy.

SPA is a forward variable selection method, which selects a wavelength at the beginning, calculates the projection value of the remaining wavelength, cycles forward, selects the wavelength corresponding to the maximum projection value, and then combines the projection vector with the wavelength until the end of the cycle (Araújo et al., 2001). SPA can minimize the collinearity between variables, extract the minimum redundant information variable group, and reduce the number of variables required to establish the model, thus improving the efficiency and speed of modeling.

CARS is a feature variable selection method that combines Monte Carlo sampling with PLS model regression coefficient (Li et al., 2009). The primary selection of the feature variables is conducted by combining the PLS regression coefficient with

exponential decay through adaptive weighted sampling each time. Then, a new PLS model is constructed based on the new subset established with the points of larger absolute weight of regression coefficient retained and the points with smaller weight removed in the PLS model. After multiple calculations, the wavelength in the subset with the smallest root mean square error of the PLS model is selected as the characteristic wavelength.

RF is a very effective algorithm, which is similar to Reversible Jump Markov Chain Monte Carlo, proposed by Li for variable selection of high-dimensional data (Li et al., 2012). It functions in an iterative manner, calculating the probability of each variable being selected in each iteration. The higher the probability, the greater the importance of the variable. The variable with the higher probability is preferred as the characteristic variable.

2.4.4 Model construction and assessment

In this study, PLS and least squares-support vector machine (LS-SVM) are selected for constructing models. LS-SVM (Suykens and Vandewalle, 1999), a machine learning algorithm based on support vector machine, is selected for constructing a regression model by adopting partial least squares linear system as loss function through nonlinear mapping function. With input variables projected into a high-dimensional feature space, and then the optimization problem converted into equality constraint conditions, this model has good generalization performance and nonlinear regression processing performance. When LS-SVM is used for analysis, appropriate kernel functions must be decided. In this study, RBF kernel function is adopted, and two parameters of the kernel function, γ and σ2, are selected by grid searching based on cross-validation. PLS (Mehmood et al., 2012), a multivariate statistical analysis method on the basis of principal component analysis, reduces the dimension by projecting independent variables and dependent variables into a new low-dimensional space, thus being capable of being used to treat the linear relationship between multiple independent variables and one or more dependent variables in a high-dimensional data set.

The evaluation metrics of the model are determination coefficient (R^2), root mean square error (RMSE), and relative percent deviation (RPD). R^2 reflects the stability of the model. The closer R^2 is to 1, the better the stability of the model is and the higher the degree of fitting is. RMSE is used for testing the predictive power of the model. The smaller the RMSE is, the better the predictive power of the model is. RPD is the ratio of sample standard deviation to RMSE. When RPD is less than 1.4, the model fails to predict the sample. When $1.4 \leq RPD < 2$, the model is considered to be of average effect and can be used for rough assessment of the samples. When $RPD \geq 2$, the model is considered to be of excellent predictive power (Khoshnoudi-Nia and Moosavi-Nasab, 2019).

The data processing process is shown in Figure 3. The calibration of the hyperspectral images and all the data processing are completed on MATLAB 2022a by encoding.

3 Results and analyses

3.1 Protein and spectral characteristics of mulberry leaves

The spectral reflection curve is drawn with the samples divided into 3 groups according to the level of protein content (Figure 4).

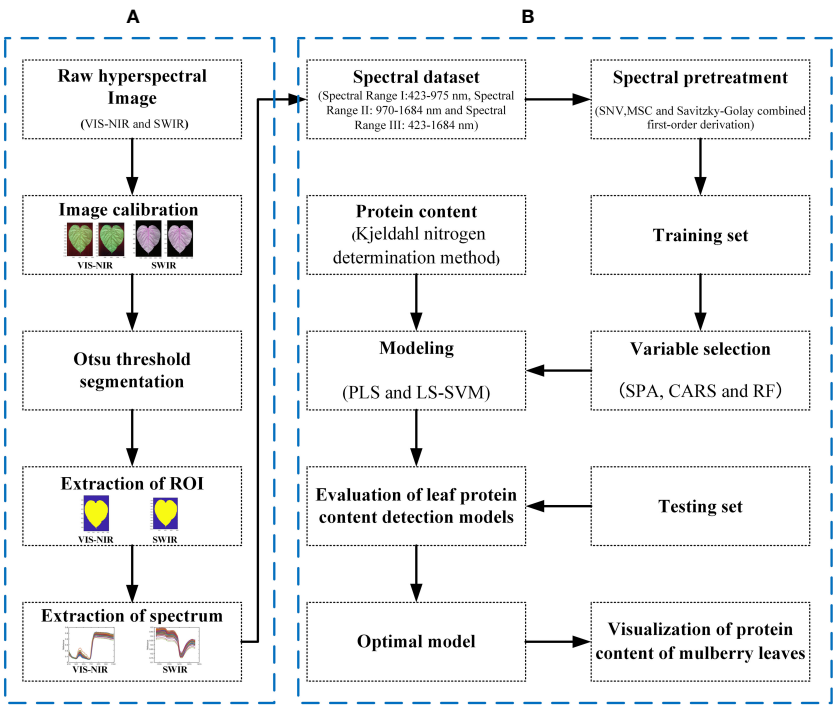


FIGURE 3
Workflow of data processing. **(A)** The raw hyperspectral image preprocessing and segmentation procedure. **(B)** The spectral processing, variable selection, and the modeling procedure.

The spectral reflection curves of mulberry leaf samples of different protein levels show the characteristics of the typical reflection spectral curve of a green plant, as follows: green peak (530–580 nm), red valley (590–670 nm), red edge (680–780 nm), high reflective platform (750–1,300 nm) related to leaf tissue structures, and the water absorption peak (1,450 nm) (Gates et al., 1965; Gausman and Allen, 1973; Gitelson et al., 1996; Tang et al., 2005). Absorption peaks of protein-associated N-H bonds have been reported at 1,020 nm and 1,510 nm in the SWIR region, but this absorption peak is not directly shown from the spectrum of Figure 4, which may be due to the fact that the absorption bands in the NIR region tend to be wide and often overlap (Curran, 1989; ElMasry et al., 2011). In addition, from Figure 4, we can see that the higher the protein content of mulberry leaves is, the lower the corresponding spectral reflectance is.

3.2 Results of feature wavelength extraction

In this study, SPA, RF, and CARS are being used individually to extract characteristic wavelengths from spectral data within three band ranges (Figure 5). In this study, the subset of bands with the smallest root Mean Square Error of Cross-Validation (RMSECV) value was selected as the characteristic band determined in the CARS and SPA algorithms. The CARS algorithm was iterated 1,000 times to ensure a comprehensive exploration. Similarly, the RF algorithm was also iterated 1,000 times to thoroughly explore the entire dataset, and by selecting the top 10 wavelength variables with a high average probability from these 1,000 runs, we obtained the characteristic wavelengths.

Analysis of the feature wavelengths extracted by using SPA, RF, and CARS shows that there are differences in the positions and numbers of the obtained feature wavelengths extracted from the spectral data undergoing the same pretreatment by using the different feature screening methods, but the extracted wavelength positions tend to concentrate in some specific bands. There are also differences in the positions and numbers of the obtained feature wavelengths extracted from the spectral data undergoing different pretreatments by using the same feature screening method, but the extracted wavelength positions tend to concentrate in some specific bands. The obtained feature wavelengths extracted in Spectral Range I mainly concentrate in 450–700 nm and 900–1,000 nm. The obtained feature wavelengths extracted in Spectral Range II mainly concentrate in 1,170–1,350 nm. The obtained feature wavelengths extracted in Spectral Range III bands mainly concentrate in the visible light region of 450–700 nm, near 800 nm, at 950 nm, and in 1,500– 1,650 nm. It is found, that the characteristic bands of proteins obtained in relevant studies are highly overlapping in the positions with the characteristic wavelengths obtained in this study. However, there are obvious differences in the specific positions and numbers. This is speculated to be caused by the heterogeneity of protein composition among different species (Caporaso et al., 2018; Ma J. et al., 2019; Ma et al., 2021; Cruz-Tirado et al., 2023). These results demonstrate the effectiveness of the applied feature screening methods (SPA, RF, and CARS) in extracting relevant wavelengths for protein content detection in mulberry leaves using HSI.

3.3 Results of modeling

Prediction models for protein content is constructed on the basis of PLS and LS-SVM, respectively, by combining three pretreatment methods, three feature wavelength screening methods, and full-band wavelength (Tables 1–3). In this study, the R^2, RMSE, and RPD of the test set are used to evaluate the predictive ability of the model, and the most suitable model for mulberry leaf content detection is selected by combining the number of variables and the predictive ability of the model.

It can be seen in Table 1, in Spectral Range I, the results obtained using the PLS models are better than those obtained using the LS-SVM models, and the S.G. first-order derivation + SPA + PLS model achieves the best performance, with an R^2 of 0.90, an RMESP of 0.85, and an RPD of 2.91. It can be seen in Table 2, in Spectral Range II, the results obtained using LS-SVM and PLS models are not much different, and the SNV + SPA + LS-SVM model achieves the best performance, with an R^2 of 0.93, an RMESP of 0.71, and an RPD of 3.83. It can be seen in Table 3, in Spectral Range III, the results obtained by using the PLS model are better than those obtained by using the LS-SVM model, and the SNV + SPA +L S-SVM model achieves the best performance, with an R^2 of 0.93, an RMESP of 0.73, and an RPD of 3.51. The above results show that there is no specific pretreatment method, characteristic

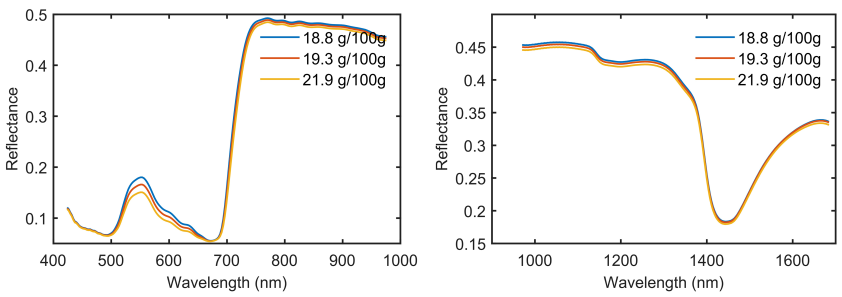


FIGURE 4
The average spectra of mulberry leaves with different protein content.

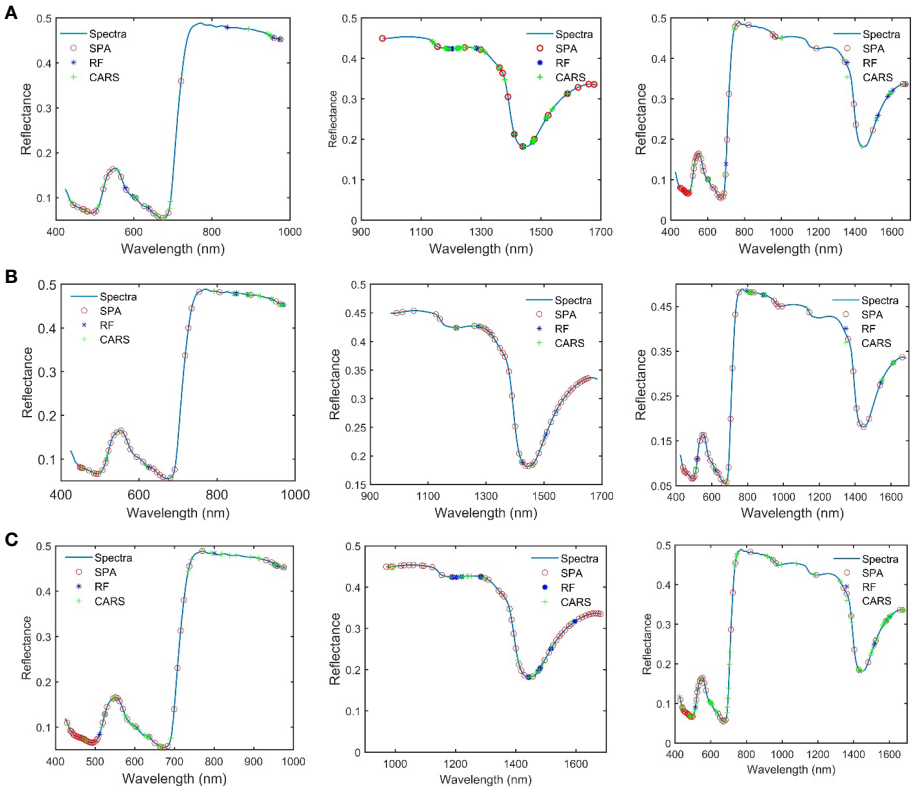


FIGURE 5
The characteristic wavelength obtained after the combination of different pretreatment and variable extraction methods. **(A)** SNV preprocessing. **(B)** Savitzky–Golay combined first-order derivation preprocessing. **(C)** MSC preprocessing.

wavelength, or modeling method optimal for all types of spectral data, and it is necessary to explore the effects of different algorithm combinations on model performance, so as to select the optimal processing method in light of different situations. LS-SVM is superior to linear methods in solving nonlinear problems, but it is sensitive to noise and error in the input data, while in Spectral Range I and III, the data used for modeling may be of poor quality or poorly correlated with the detection of protein content spectral data, so PLS algorithms obtain better results than LS-SVM algorithms (Suykens et al., 2002; Wang and Hu, 2005).

As can be seen from Tables 1–3, better results are achieved by using the models based on the feature extraction method, compared with the full-band models, which is due to the fact that the full-band spectral data have some redundant and interference information, and this is an indication that the feature extraction method can effectively reduce the redundant information between adjacent spectral bands and improve the accuracy of models. Two hyperspectral cameras with different wavelength ranges were compared, and the overall performance of the predictive model developed in the SWIR region shows better predictive power and robustness than that established in the Vis-NIR region, which is exactly opposite to the results of Ma et al. (Ma J. et al., 2019). They obtained better results in detecting pork protein by using spectral data of the Vis-NIR region. However, in many other protein detection studies, good prediction results are obtained by using spectral data of the SWIR region (Talens et al., 2013; Ma et al., 2021; He et al., 2023). In this study, compared with the models constructed based on Spectral Range II spectral data, the model based on Spectral Range III spectral data fails to show better accuracy, although it obtains richer spectral information. This may result from the spectrum of the Spectral Range III region containing more redundant information related to the detection of protein content of mulberry leaves. The above results show that the SWIR region is the optimal spectral range for mulberry leaf protein prediction.

Previous studies have explored the feasibility of HSI for the non-destructive detection of protein content; however, few studies have attempted to determine the optimal spectral range for measuring proteins, especially for fresh mulberry leaves. In this study, the best results are obtained by combining the SWIR HSI acquisition system based on InGaAs detectors with SNV + SPA + LS-SVM, with an R^2 of the test set of up to 0.93, an RMSE of only 0.71 g/100 g, and an RPD of up to 3.83. The results show that the model is qualified for detecting and analyzing the protein content of mulberry leaves.

3.4 Visualization of protein content of mulberry leaves

The distribution of protein content in mulberry leaves has not been reported. In the practical application of non-destructive

10.3389/fpls.2023.1275004

TABLE 1 Results of models in spectral range I.

Pretreatment	Variable selection	No. of variables	PLS				LS-SVM		
			PCs	R^2	RMSE	RPD	R^2	RMSE	RPD
SNV	Full bands	204	17	0.82	1.14	2.28	0.60	1.71	1.60
	SPA	24	21	0.89	0.91	2.62	0.86	1.01	2.69
	CARS	23	18	0.88	0.93	2.68	0.84	1.09	2.53
	RF	10	9	0.88	0.92	2.81	0.84	1.08	2.49
S.G. first-order derivation	Full bands	204	16	0.90	0.85	2.94	0.70	1.47	1.85
	SPA	35	23	0.90	0.85	2.91	0.82	1.15	2.36
	CARS	26	11	0.66	1.57	1.63	0.69	1.51	1.80
	RF	10	7	0.69	1.49	1.61	0.72	1.43	1.91
MSC	Full bands	204	16	0.86	1.01	2.59	0.60	1.71	1.59
	SPA	51	18	0.90	0.87	2.87	0.76	1.31	2.08
	CARS	39	14	0.87	0.96	1.96	0.73	1.40	1.95
	RF	10	8	0.89	0.90	3.20	0.85	1.04	2.63

detection technology for mulberry leaf protein, the visualization of the protein content of mulberry leaves can not only provide valuable insights for merchants to classify the freshness and quality of mulberry leaves more intuitively, but also aid researchers in conducting plant physiology studies related to mulberry leaves. By extracting spectral data from all pixels of the leaves, a distribution map is generated by using the established SNV + SPA + LS-SVM model to visualize the spatial distribution of protein content in mulberry leaves. The level of protein content is represented by the depth of shade, as depicted in Figure 6. It should be noted that the variety, harvest time, and maturity significantly influence the nutrient content of mulberry leaves. Previous studies have indicated that the protein content of mulberry leaves decreases with increasing ripeness (Ramesh et al., 2021). In the visualization results of this study, it can be seen that tender leaves exhibit higher protein content compared to mature leaves, which is consistent with the above findings. Additionally, the visualized distribution of the protein content of mulberry leaves shows that the protein of healthy mulberry leaves is essentially evenly distributed in the mesophyll, while the protein content in the veins is extremely low. This is due to heterogeneity, and the fact that the protein content varies across different locations within the sample and the leaf vein is mainly composed of cellulose and conductive substances with no capacity of storing energy (Fukuda, 2004; Jiang et al., 2022).

TABLE 2 Results of models in spectral range II.

Pretreatment	Variable selection	No. of variables	PLS				LS-SVM		
			PCs	R^2	RMSE	RPD	R^2	RMSE	RPD
SNV	Full bands	204	15	0.86	1.02	2.17	0.87	0.97	2.81
	SPA	15	15	0.92	0.78	3.07	0.93	0.71	3.83
	CARS	26	9	0.85	1.05	1.95	0.86	1.03	2.24
	RF	10	8	0.81	1.19	2.15	0.80	1.21	2.65
S.G. first-order derivation	Full bands	204	14	0.85	1.04	2.13	0.89	0.90	3.04
	SPA	42	19	0.86	1.01	2.25	0.91	0.80	3.40
	CARS	13	7	0.85	1.05	2.12	0.88	0.94	2.91
	RF	10	7	0.80	1.20	1.84	0.86	1.01	2.70
MSC	Full bands	204	14	0.88	0.94	2.36	0.87	0.98	2.78
	SPA	45	15	0.92	0.78	3.09	0.86	1.02	2.68
	CARS	26	8	0.81	1.17	1.93	0.81	1.17	2.34
	RF	10	9	0.74	1.37	1.48	0.71	1.44	1.89

10.3389/fpls.2023.1275004

TABLE 3 Results of models in spectral range III.

Pretreatment	Variable selection	No. of variables	PLS				LS-SVM		
			PCs	R^2	RMSE	RPD	R^2	RMSE	RPD
SNV	Full bands	405	18	0.90	0.84	3.04	0.88	0.93	2.93
	SPA	50	21	0.92	0.78	3.20	0.92	0.78	3.51
	CARS	19	14	0.75	1.36	1.40	0.70	1.48	1.30
	RF	10	10	0.67	1.54	1.78	0.39	2.10	1.84
S.G. first-order derivation	Full bands	204	22	0.93	0.72	3.39	0.91	0.82	3.31
	SPA	43	28	0.92	0.77	3.20	0.90	0.85	3.22
	CARS	19	7	0.87	0.97	2.49	0.85	1.04	2.62
	RF	10	6	0.88	0.95	2.61	0.88	0.92	2.97
MSC	Full bands	204	16	0.90	0.86	3.00	0.87	0.97	2.80
	SPA	48	19	0.93	0.73	3.51	0.86	0.99	2.75
	CARS	77	16	0.92	0.76	3.39	0.87	0.99	2.75
	RF	10	9	0.79	1.24	1.79	0.29	2.28	1.19

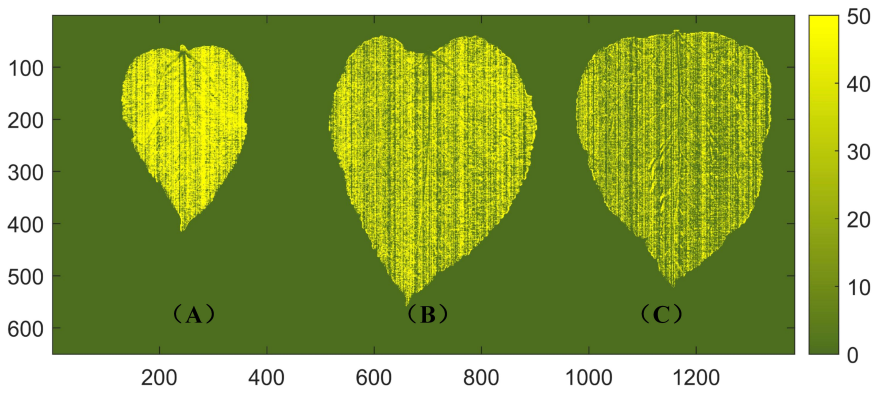


FIGURE 6
Visualization of protein content in mulberry leaves. **(A)** A young leaf with a protein content of 45.7 g/100 g. **(B)** A middle mature leaf with a protein content of 26.3 g/100 g. **(C)** A mature leaf with a protein content of 16.3 g/100 g.

4 Conclusion

The protein content of mulberry leaves is a crucial indicator for assessing their quality. In this study, we aimed to develop a rapid and non-destructive method for detecting the protein content of mulberry leaves using HSI technology. The feasibility of using HSI technology within the spectral range of 400–1,000 nm and 900–1,700 nm for non-destructive detection of mulberry leaf protein content is investigated. By comparing different spectral ranges of the HSI acquisition system and utilizing various data processing methods, including preprocessing, variable extraction, and modeling, prediction models for protein content detection are constructed. The results demonstrated that the best performance was achieved by combining the spectral data from 900–1,700 nm with SNV + SPA + LS-SVM. This approach yielded a testing set R^2 value of up to 0.93, an RMSE of only 0.71 g/100 g, and an RPD of up to 3.83. Furthermore, the visualization of the protein content distribution in mulberry leaves based on the best model revealed that healthy leaves exhibited an even distribution of protein content throughout the mesophyll, with lower protein concentrations observed in the leaf veins.

These findings show the optimal spectral range for mulberry leaf protein prediction and highlight the potential of utilizing SWIR HSI combined with the SNV–SPA–LS-SVM algorithm for rapid, non-destructive, and high-precision detection of protein content in mulberry leaves. The developed method can provide valuable insights for assessing the quality of mulberry leaves in a non-invasive manner, enabling efficient monitoring and optimization of mulberry leaf quality.

Data availability statement

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

Author contributions

XL: Conceptualization, Software, Writing – original draft. FP: Data curation, Investigation, Writing – original draft. ZW: Investigation, Writing – original draft. GH: Conceptualization, Funding acquisition, Supervision, Writing – review & editing. JL: Methodology, Supervision, Writing – review & editing.

Funding

The authors declare financial support was received for the research, authorship, and/or publication of this article. This research was funded by the Natural Science Foundation of Chongqing, China (Grant Nos. CSTB2023NSCQ-MSX1018 and CSTB2023NSCQ- MSX0043), Municipal financial scientific research project of Chongqing Academy of Agricultural Sciences (Grant No. cqaas2023sjczqn001), the Chongqing Performance Incentive and Guidance Project for Scientific Research Institutions (Grant Nos. cstc2020jxjl80008 and cstc2022jxjl80021), and the Excellent Germplasm Innovation Project of Chongqing (Grant No. NKY2021AA016).

Conflict of interest

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

Publisher's note

References

Araújo, M. C. U., Saldanha, T. C. B., Galvao, R. K. H., Yoneyama, T., Chame, H. C., and Visani, V. (2001). The successive projections algorithm for variable selection in spectroscopic multicomponent analysis. *Chemom. Intell. Lab. Syst.* 57 (2), 65–73. doi: 10.1016/S0169-7439(01)00119-8

Barnes, R., Dhanoa, M. S., and Lister, S. J. (1989). Standard normal variate transformation and de-trending of near-infrared diffuse reflectance spectra. *Appl. Spectrosc.* 43 (5), 772–777. doi: 10.1366/0003702894202201

Caporaso, N., Whitworth, M. B., and Fisk, I. D. (2018). Protein content prediction in single wheat kernels using hyperspectral imaging. *Food Chem.* 240, 32–42. doi: 10.1016/j.foodchem.2017.07.048

Chelladurai, V., and Jayas, D. (2014). "Near-infrared imaging and spectroscopy," in *Imaging with Electromagnetic Spectrum: Applications in Food and Agriculture* (Berlin, Heidelberg: Springer), 87–127.

Cheng, J.-H., Jin, H., Xu, Z., and Zheng, F. (2017). NIR hyperspectral imaging with multivariate analysis for measurement of oil and protein contents in peanut varieties. *Anal. Methods* 9 (43), 6148–6154. doi: 10.1039/C7AY02115A

Chromý, V., Vinklárková, B., Šprongl, L., and Bittová, M. (2015). The Kjeldahl method as a primary reference procedure for total protein in certified reference materials used in clinical chemistry. I. A review of Kjeldahl methods adopted by laboratory medicine. *Crit. Rev. Anal. Chem.* 45 (2), 106–111. doi: 10.1080/10408347.2014.892820

Cruz-Tirado, J. P., Amigo, J. M., and Barbin, D. F. (2023). Determination of protein content in single black fly soldier (Hermetia illucens L.) larvae by near infrared hyperspectral imaging (NIR-HSI) and chemometrics. *Food Control* 143, 109266. doi: 10.1016/j.foodcont.2022.109266

Curran, P. J. (1989). Remote sensing of foliar chemistry. *Remote Sens. Environ.* 30 (3), 271–278. doi: 10.1016/0034-4257(89)90069-2

Denholm, K., Haggerty, A., Mason, C., and Ellis, K. (2021). Comparison of tests for failure of passive transfer in neonatal calf serum using total protein refractometry and the biuret method. *Prev. Vet. Med.* 189, 105290. doi: 10.1016/j.prevetmed.2021.105290

ElMasry, G., Sun, D.-W., and Allen, P. (2011). Non-destructive determination of water-holding capacity in fresh beef by using NIR hyperspectral imaging. *Food Res. Int.* 44 (9), 2624–2633. doi: 10.1016/j.foodres.2011.05.001

ElMasry, G., Sun, D.-W., and Allen, P. (2013). Chemical-free assessment and mapping of major constituents in beef using hyperspectral imaging. *J. Food Eng.* 117 (2), 235–246. doi: 10.1016/j.jfoodeng.2013.02.016

Fukuda, H. (2004). Signals that control plant vascular cell differentiation. Nature Reviews Molecular. *Cell Biol.* 5 (5), 379–391. doi: 10.1038/nrm1364

Gates, D. M., Keegan, H. J., Schleter, J. C., and Weidner, V. R. (1965). Spectral properties of plants. *Appl. Optics.* 4 (1), 11–20. doi: 10.1364/AO.4.000011

Gausman, H., and Allen, W. (1973). Optical parameters of leaves of 30 plant species. *Plant Physiol. Biochem.* 52 (1), 57–62. doi: 10.1104/pp.52.1.57

Gitelson, A. A., Merzlyak, M. N., and Lichtenthaler, H. K. (1996). Detection of red edge position and chlorophyll content by reflectance measurements near 700 nm. *J. Plant Physiol.* 148 (3), 501–508. doi: 10.1016/S0176-1617(96)80285-9

Gryn-Rynko, A., Bazylak, G., and Olszewska-Slonina, D. (2016). New potential phytotherapeutics obtained from white mulberry (Morus alba L.) leaves. *Biomed. Pharmacother.* 84, 628–636. doi: 10.1016/j.biopha.2016.09.081

He, H.-J., Wang, Y., Wang, Y., Liu, H., Zhang, M., and Ou, X. (2023). Simultaneous quantifying and visualizing moisture, ash and protein distribution in sweet potato [Ipomoea batatas (L.) Lam] by NIR hyperspectral imaging. *Food Chem.: X* 18, 100631.

Isaksson, T., and Næs, T. (1988). The effect of multiplicative scatter correction (MSC) and linearity improvement in NIR spectroscopy. *Appl. Spectrosc.* 42 (7), 1273–1284. doi: 10.1366/0003702884429869

Jiang, X., Tian, J., Huang, H., Hu, X., Han, L., Huang, D., et al. (2022). Nondestructive visualization and quantification of total acid and reducing sugar contents in fermented grains by combining spectral and color data through hyperspectral imaging. *Food Chem.* 386, 132779. doi: 10.1016/j.foodchem.2022.132779

Jin, X., Xiao, Z.-Y., Xiao, D.-X., Dong, A., Nie, Q.-X., Wang, Y.-N., et al. (2022). Quantitative inversion model of protein and fat content in milk based on hyperspectral techniques. *Int. Dairy J.* 134, 105467. doi: 10.1016/j.idairyj.2022.105467

Khoshnoudi-Nia, S., and Moosavi-Nasab, M. (2019). Prediction of various freshness indicators in fish fillets by one multispectral imaging system. *Sci. Rep.* 9 (1), 14704.

Ledoux, M., and Lamy, F. (1986). Determination of proteins and sulfobetaine with the folin-phenol reagent. *Anal. Biochem.* 157 (1), 28–31. doi: 10.1016/0003-2697(86)90191-0

Li, H., Liang, Y., Xu, Q., and Cao, D. (2009). Key wavelengths screening using competitive adaptive reweighted sampling method for multivariate calibration. *Anal. Chim. Acta* 648 (1), 77–84. doi: 10.1016/j.aca.2009.06.046

Li, Q., Liu, F., Liu, J., Liao, S., and Zou, Y. (2019). Mulberry leaf polyphenols and fiber induce synergistic antiobesity and display a modulation effect on gut microbiota and metabolites. *Nutrients* 11 (5), 1017. doi: 10.3390/nu11051017

Li, H.-D., Xu, Q.-S., and Liang, Y.-Z. (2012). Random frog: an efficient reversible jump Markov Chain Monte Carlo-like approach for variable selection with applications to gene selection and disease classification. *Anal. Chim. Acta.* 740, 20–26. doi: 10.1016/j.aca.2012.06.031

Ma, J., Cheng, J.-H., Sun, D.-W., and Liu, D. (2019). Mapping changes in sarcoplasmatic and myofibrillar proteins in boiled pork using hyperspectral imaging with spectral processing methods. *LWT* 110, 338–345. doi: 10.1016/j.lwt.2019.04.095

Ma, C., Ren, Z., Zhang, Z., Du, J., Jin, C., and Yin, X. (2021). Development of simplified models for nondestructive testing of rice (with husk) protein content using hyperspectral imaging technology. *Vib. Spectrosc.* 114, 103230. doi: 10.1016/j.vibspec.2021.103230

Ma, Y., Zhang, G.-Z., and Rita-Cindy, S. A.-A. (2019). Quantification of water, protein and soluble sugar in mulberry leaves using a handheld near-infrared spectrometer and multivariate analysis. *Molecules* 24 (24), 4439. doi: 10.3390/molecules24244439

Mehmood, T., Liland, K. H., Snipen, L., and Sæbo, S. (2012). A review of variable selection methods in partial least squares regression. *Chemom. Intell. Lab. Syst.* 118, 62–69. doi: 10.1016/j.chemolab.2012.07.010

Onoyama, H., Ryu, C., Suguri, M., and Iida, M. (2018). Estimation of rice protein content before harvest using ground-based hyperspectral imaging and region of interest analysis. *Precis. Agric.* 19 (4), 721–734. doi: 10.1007/s11119-017-9552-3

Pu, H., Sun, D.-W., Ma, J., Liu, D., and Kamruzzaman, M. (2014). Hierarchical variable selection for predicting chemical constituents in lamb meats using hyperspectral imaging. *J. Food Eng.* 143, 44–52. doi: 10.1016/j.jfoodeng.2014.06.025

Ramesh, H. L., Shivashankarappa, L. H., and Yoganandamurthy, V. N. (2021). Gamma irradiation enhanced leaf bioactive components and bioassay parameters in M5 mulberry (Morus sp.) mutant. *Am. J. Plant Sci.* 12 (12), 1945–1962. doi: 10.4236/ajps.2021.1212134

Riche, D. M., Riche, K. D., East, H. E., Barrett, E. K., and May, W. L. (2017). Impact of mulberry leaf extract on type 2 diabetes (Mul-DM): a randomized, placebo-controlled pilot study. *Complement. Ther. Med.* 32, 105–108. doi: 10.1016/j.ctim.2017.04.006

Savitzky, A., and Golay, M. J. (1964). Smoothing and differentiation of data by simplified least squares procedures. *Anal. Chem.* 36 (8), 1627–1639. doi: 10.1021/ac60214a047

Shenk, J. S., Workman Jr., J. J., and Westerhaus, M. O. (2007). *"Application of NIR spectroscopy to agricultural products," in Handbook of near-infrared analysis"* CRC Press, 365–404.

Srivastava, S., Kapoor, R., Thathola, A., Srivastava, R., and Nutrition, (2003). Mulberry (Moms alba) leaves as human food: a new dimension of sericulture. *Int. J. Food Sci.* 54 (6), 411–416. doi: 10.1080/0963748310001622288

Sun, C., Wu, W., Ma, Y., Min, T., Lai, F., and Wu, H. (2017). Physicochemical, functional properties, and antioxidant activities of protein fractions obtained from mulberry (morus atropurpurea roxb.) leaf. *Int. J. Food Prop.* 20 (sup3), S3311–S3325. doi: 10.1080/10942912.2016.1238389

Suykens, J. A., De Brabanter, J., Lukas, L., and Vandewalle, J. (2002). Weighted least squares support vector machines: robustness and sparse approximation. *Neurocomputing* 48 (1-4), 85–105. doi: 10.1016/S0925-2312(01)00644-0

Suykens, J. A., and Vandewalle, J. (1999). Least squares support vector machine classifiers. *Neural Process. Lett.* 9, 293–300. doi: 10.1023/A:1018628609742

Talens, P., Mora, L., Morsy, N., Barbin, D. F., ElMasry, G., and Sun, D.-W. (2013). Prediction of water and protein contents and quality classification of Spanish cooked ham using NIR hyperspectral imaging. *J. Food Eng.* 117 (3), 272–280. doi: 10.1016/j.jfoodeng.2013.03.014

Tang, Y., Wen, X., and Lu, C. (2005). Differential changes in degradation of chlorophyll–protein complexes of photosystem I and photosystem II during flag leaf senescence of rice. *Plant Physiol. Biochem.* 43 (2), 193–201. doi: 10.1016/j.plaphy.2004.12.009

Thabti, I., Elfalleh, W., Tlili, N., Ziadi, M., Campos, M. G., and Ferchichi, A. (2014). Phenols, flavonoids, and antioxidant and antibacterial activity of leaves and stem bark of Morus species. *Int. J. Food Prop.* 17 (4), 842–854. doi: 10.1080/10942912.2012.660722

Wang, H., and Hu, D. "Comparison of SVM and LS-SVM for Regression," in *2005 International conference on neural networks and brain*. 279–283.

Yu, Y., Li, H., Zhang, B., Wang, J., Shi, X., Huang, J., et al. (2018). Nutritional and functional components of mulberry leaves from different varieties: Evaluation of their potential as food materials. *Int. J. Food Prop.* 21 (1), 1495–1507. doi: 10.1080/10942912.2018.1489833

Zhang, C., Liu, F., Kong, W., and He, Y. (2015). Application of visible and near-infrared hyperspectral imaging to determine soluble protein content in oilseed rape leaves. *Sensors.* 15 (7), 16576–16588. doi: 10.3390/s150716576

frontiers | Frontiers in Plant Science

TYPE Original Research
PUBLISHED 08 November 2023
DOI 10.3389/fpls.2023.1283230

OPEN ACCESS

EDITED BY
Qiang Lyu,
Southwest University, China

REVIEWED BY
Yunchao Tang,
Guangxi University, China
Liantao Liu,
Hebei Agricultural University, China

*CORRESPONDENCE
Shengping Lv
lvshengping@scau.edu.cn

RECEIVED 25 August 2023
ACCEPTED 20 October 2023
PUBLISHED 08 November 2023

CITATION
Li X, Zhang Z, Lv S, Liang T, Zou J, Ning T and Jiang C (2023) Detection of breakage and impurity ratios for raw sugarcane based on estimation model and MDSC-DeepLabv3+. *Front. Plant Sci.* 14:1283230. doi: 10.3389/fpls.2023.1283230

Detection of breakage and impurity ratios for raw sugarcane based on estimation model and MDSC-DeepLabv3+

Xin Li, Zhigang Zhang, Shengping Lv*, Tairan Liang, Jianmin Zou, Taotao Ning and Chunyu Jiang

College of Engineering, South China Agricultural University, Guangzhou, China

Broken cane and impurities such as top, leaf in harvested raw sugarcane significantly influence the yield of the sugar manufacturing process. It is crucial to determine the breakage and impurity ratios for assessing the quality and price of raw sugarcane in sugar refineries. However, the traditional manual sampling approach for detecting breakage and impurity ratios suffers from subjectivity, low efficiency, and result discrepancies. To address this problem, a novel approach combining an estimation model and semantic segmentation method for breakage and impurity ratios detection was developed. A machine vision-based image acquisition platform was designed, and custom image and mass datasets of cane, broken cane, top, and leaf were created. For cane, broken cane, top, and leaf, normal fitting of mean surface densities based on pixel information and measured mass was conducted. An estimation model for the mass of each class and the breakage and impurity ratios was established using the mean surface density and pixels. Furthermore, the MDSC-DeepLabv3+ model was developed to accurately and efficiently segment pixels of the four classes of objects. This model integrates improved MobileNetv2, atrous spatial pyramid pooling with deepwise separable convolution and strip pooling module, and coordinate attention mechanism to achieve high segmentation accuracy, deployability, and efficiency simultaneously. Experimental results based on the custom image and mass datasets showed that the estimation model achieved high accuracy for breakage and impurity ratios between estimated and measured value with R^2 values of 0.976 and 0.968, respectively. MDSC-DeepLabv3+ outperformed the compared models with mPA and mIoU of 97.55% and 94.84%, respectively. Compared to the baseline DeepLabv3+, MDSC-DeepLabv3+ demonstrated significant improvements in mPA and mIoU and reduced Params, FLOPs, and inference time, making it suitable for deployment on edge devices and real-time inference. The average relative errors of breakage and impurity ratios between estimated and measured values were 11.3% and 6.5%, respectively. Overall, this novel approach enables high-precision, efficient, and intelligent detection of breakage and impurity ratios for raw sugarcane.

KEYWORDS

raw sugarcane, breakage ratio, impurity ratio, estimation model, MDSC-DeepLabv3+

1 Introduction

Sugarcane is an important raw material for the sugar industry worldwide. In China, sugarcane-based sugar production reached 4.6 million tons in 2022, which is 4.3 times that of beet sugar (National Development and Reform Commission, 2023). In recent years, the use of machine-harvested sugarcane has been steadily increasing, with plans to reach 30% of total sugarcane harvest in China by 2025 (Chinese government website, 2018). Machine harvesting significantly improves efficiency and reduces labor intensity; however, it also leads to higher ratios of broken cane and impurities such as top, leaf, which can negatively impact the yield of the sugar manufacturing process. As a result, the breakage and impurity ratios are crucial indicators for assessing the quality and pricing of raw sugarcane in practice, and determining these two ratios is indispensable for sugar refineries. Unfortunately, the commonly used manual sampling approach for detecting breakage and impurity ratios brings several issues, including strong subjectivity, low efficiency, and significant result discrepancies.

To address the aforementioned problem, an estimation model was established, and machine vision technology was employed to provide a more objective, efficient, accurate, and intelligent approach for quantifying the cane, broken cane, and impurities, as well as the ratios of breakage and impurity. This enables seamless integration with the sugarcane harvesting and sugar processing stages. Both cane and broken cane can be used as raw materials, but broken cane is considered in mass deduction by sugar refineries because it results in the loss of sugar content and impacts the quality of the final sugar product. The sugarcane top, leaf, root, sand, gravel, and soil and so forth are collectively referred to as impurities (Guedes and Pereira, 2018). Adjusting the height between the harvester's cutting device and the ridge surface will reduce the introduction of sand, gravel, and soil during sugarcane harvesting. Furthermore, when the mechanical harvester operates smoothly and adheres to specifications, it noticeably decreases the levels of mud, stone, and cane root (Xie et al., 2018). Mechanical removal methods, such as vibration, can often be used to screen out the sand, gravel, and soil (Martins and Ruiz, 2020). However, the top, leaf and cane root are unavoidable impurities as they are naturally part of each sugarcane stem (de Mello et al., 2022). Regarding cane root, object detection can be utilized to count its quantities. Combining this with the average weight of the cane root helps predict the mass of root impurity after excluding sand, gravel and soil. Based on the quality detection practice of sugar refineries, the four categories of cane, broken cane, top, and leaf are selected as the detection objects in this study.

Estimation models and machine vision technology have been widely used for the detection and monitoring of impurities in grain crops such as rice, wheat, and corn. For example, Chen et al. (2020) used morphological features and a decision tree for the classification of rice grains and impurities with 76% accuracy to optimize combine harvester parameters. Liu et al. (2023) proposed a NAM-EfficientNetv2 lightweight segmentation approach for rapid online detection of rice seed and impurities in harvesters, achieving high evaluation index F1 scores of 95.26% and 93.27% for rice grain and impurities, respectively. To improve accuracy in wheat and impurity recognition, Shen et al. (2019) constructed a dataset and trained a recognition model called WheNet based on Inception_v3, achieving a recall rate of 98% and an efficiency of 100ms per image. Chen et al. (2022) designed a vision system based on DeepLabv3+ to identify seeds and impurities in wheat, obtaining mean pixel accuracy (mPA) values of 86.86% and 89.91% for grains and impurities, and mean intersection over union (mIoU) scores of 0.7186 and 0.7457, respectively. For the detection of impurities in the corn deep-bed drying process, Li et al. (2022) employed a multi-scale color recovery algorithm to enhance images and eliminate noise. They used HSV color space parameter thresholds and morphological operations for segmentation and achieved F1 scores of 83.05%, 83.87%, and 87.43% for identifying broken corncob, broken bract, and crushed stone, respectively. Liu et al. (2022) developed a CPU-Net semantic segmentation model based on U-Net, incorporating the convolutional block attention module (CBAM) and pyramid pooling modules to improve segmentation accuracy for monitoring corn kernels and their impurities. They established a mass-pixel linear regression model to calculate the kernel impurity rate and experimental results demonstrated that CPU-Net outperforms other comparative approaches with average mIoU, mPA, and inference time scores of 97.31%, 98.71%, and 158.4ms per image, respectively. The average relative error between the impurity rate obtained by the model and manual statistics was 4.64%.

Detection of impurities in cash crops such as soybean, cotton, and walnut during harvesting or processing has also been extensively studied in recent years. Momin et al. (2017) used HSI to segment the image background of soybean with three categories of impurities. They employed various image processing techniques, such as median blur, morphological operations, watershed transformation, projection area-based analysis, and circle detection, for feature recognition of soybean and impurities. The experimental results showed pixel accuracy of 96%, 75%, and 98% for split bean, contaminated bean, and defective bean, and stem/pod, respectively. Jin et al. (2022) developed an improved UNet segmentation model to address issues of soybean sticking, stacking, and complex semantics in images. The experimental results demonstrated comprehensive evaluation index values of 95.50%, 91.88%, and 94.34% for complete grain, broken grain, and impurity segmentation, respectively, with a mIoU of 86.83%. The field experiment indicated mean absolute errors of 0.18 and 0.10 percentage points for fragmentation and impurity rate between the model-based value and the measured value, respectively. For real-time detection of impurity ratio in cotton processing, Zhang et al. (2022) utilized the enhanced Canny algorithm to segment cotton and its impurities. They employed YOLOv5 to identify the segmented objects and determine their respective categories. They also developed an estimation model for the impurity ratio based on segmented volume and estimated mass and utilized a multithread technique to shorten the processing time, achieving a 43.65% reduction compared to that of a single thread. To improve the recognition accuracy of white and near-cotton-colored impurities in raw cotton, Xu et al. (2023) proposed a weighted feature fusion module and a decoupled detection strategy to enhance the detection head of YOLOv4-tiny. The proposed method decreased

computation during the inference process, boosted the speed of inference, and enhanced the accuracy of cotton impurity localization. Experimental results showed a respective increase of 10.35% and 6.9% in mAP and frames per second (FPS) compared to the baseline YOLOv4-tiny. The detection accuracy of white and near cotton-colored impurities in raw cotton reached 98.78% and 98%, respectively. To achieve real-time segmentation of juglans impurity, Rong et al. (2020) proposed a hybrid approach by combining a segmentation model based on a multi-scale residual full convolutional network and a classification method based on a convolutional network. The proposed method accurately segmented 99.4% and 96.5% of the object regions in the test and validation images, respectively, with a segmentation time of within 60ms for each image. Yu L. et al. (2023) presented an improved YOLOv5 with lower parameters and quicker speed for walnut kernel impurity detection by incorporating target detection layers, CBAM, transformer-encoder, and GhostNet. The results indicated a mAP of 88.9%, which outperformed the baseline YOLOv5 by 6.7%.

In recent years, researchers have also achieved notable progress in the field of impurity detection in sugarcane. Guedes and Pereira (2019) constructed an image dataset comprising 122 different combinations of sugarcane stalk, vegetal plant part, and soil to evaluate the impurity amount. They converted color samples into color histograms with ten color scales and employed three classifiers, namely soft independent modeling of class analogy, partial least squares discriminant analysis (PLS-DA), and *k* nearest neighbors (KNN), to classify cane and its impurities. Guedes et al. (2020) further proposed an analytical method using artificial neural networks (ANNs) combined with the ten color histograms to predict the content of sugarcane in the presence of impurities. The experimental results demonstrated correlation coefficients of 0.98, 0.93, and 0.91 for the training, validation, and test sets, respectively. Aparatana et al. (2020) employed principal component analysis (PCA), PLS-DA, and support vector machine (SVM) to classify and differentiate sugarcane and impurities, including green leaf, dry leaf, stone, and soil, based on their spectral information. The research findings indicated that PCA, PLS-DA, and SVM achieved classification rates of 90%, 92.9%, and 98.2%, respectively. Dos Santos et al. (2021) used a similar mechanism by combining ten color histograms and ANNs to classify raw sugarcane. They achieved 100% accurate classification for two ranges of raw sugarcane in the samples, from 90 to 100 wt% and from 41 to 87 wt%. However, these studies mentioned above recognize raw sugarcane and impurities based on their color features, making it difficult to differentiate objects with inter-class similarity, such as sugarcane top and leaf, which have similar color features at the pixel level. Additionally, these methods may not be suitable for practical situation with multiple combinations of impurities in arbitrary proportions, which present significant challenges in building samples with a vast combination of weight percentages of impurities.

From the perspective of recognition tasks, the aforementioned studies can be categorized into three types: image classification, object detection, and semantic segmentation. Image classification-based approaches (Momin et al., 2017; Guedes and Pereira, 2019; Shen et al., 2019; Aparatana et al., 2020; Chen et al., 2020; Guedes et al., 2020; Dos Santos et al., 2021; Li et al., 2022) cannot capture pixel-level information for subsequent construction of a mass-pixel fitting model. Object detection can be utilized for real-time classification and localization of crops and impurities (Zhang et al., 2022; Xu et al., 2023; Yu J. et al., 2023), but they still cannot support subsequent mass estimation based on pixels of detected objects. Semantic segmentation, on the other hand, enables pixel-wise classification of an image and facilitates the precise determination of the number of pixels and their respective categories in a specific region. Mass-pixel fitting models can be established by combining the number of pixels and the actual mass of each category of object (Rong et al., 2020; Chen et al., 2022; Jin et al., 2022; Liu et al., 2022; Liu et al., 2023), thus supporting the quantitative analysis of the quality of the detected objects. In order to quantify the ratio of breakage and impurity in raw sugarcane, semantic segmentation technology was utilized to abstract the of raw sugarcane and impurities in this study. However, the aforementioned approaches and findings are difficult to be directly applied to the detection of sugarcane and impurities in this study. Firstly, there is currently a lack of image databases that include raw sugarcane and impurities. Secondly, the estimation models developed in the above studies are only suitable for relatively stable scenarios of surface density (mass/pixel) for each detection category. However, the surface density of broken cane varies significantly due to different degrees of breakage, and the residual leaf at the top of the cane is scattered, resulting in a more varied surface density. Therefore, it is necessary to establish a corresponding image dataset and segmentation model for the detection of raw sugarcane and impurities and build new estimation model for quality evaluation based on segmented pixels.

Popular and widely applied deep learning (DL)-based semantic segmentation approaches have achieved excellent results in image processing in agriculture (Luo et al., 2023). Among these approaches, end-to-end semantic segmentation models like FCN, UNet, PSPNet, and DeepLabv3+ have demonstrated good performance with simple structures. DeepLabv3+ in particular has gained significant popularity and has been extensively enhanced due to its exceptional segmentation accuracy, making it a widely practiced and verified model in agricultural applications. For instance, Wu et al. (2021) developed an enhanced version of DeepLabv3+ to segment abnormal leaves in hydroponic lettuce. Peng et al. (2023) constructed an RDF-DeepLabv3+ for segmenting lychee *stem*. Zhu et al. (2023) proposed a two-stage DeepLabv3+ with adaptive loss for the segmentation of apple leaf disease images in complex scenes. Wu et al. (2023) utilized Deeplabv3+ and post-processing image analysis techniques for precise segmentation and counting of banana bunches. Their findings indicated that DeepLabv3+-based segmentation models can effectively perform pixel-level segmentation of crop objects, and the segmentation effects were superior to those of compared approaches. In this study, DeepLabv3+ was adopted for the semantic segmentation of raw sugarcane and impurities, and efforts were made to further improve its segmentation accuracy, reduce parameters, and optimize inference time.

This study aims to address the detection of breakage and impurity ratios in raw sugarcane. The specific research content of this study includes: (1) Designing a machine vision-based acquisition platform for online image collection of raw sugarcane (cane, broken cane) and impurities (top, leaf). Custom datasets of masses and corresponding images were constructed. (2) Establishing a normal fitting model to determine the mean surface density of each class based on measured masses and extracted pixels. Additionally, an estimation model was developed to assess the ratios of breakage and impurity using the estimated mass of each class, along with their pixels and fitted mean surface density. (3) Developing a MDSC-DeepLabv3+ model for accurate segmentation of raw sugarcane and impurity pixels based on DeepLabv3+. The model was further improved by incorporating improved MobileNetv2, atrous spatial pyramid pooling (ASPP) with deepwise separable convolution (DSC) and strip pooling (SP) named ASPP_DS, and coordinate attention (CA) mechanism to enhance segmentation accuracy, reduce parameters, and optimize inference time. (4) Conducting experiments to verify the accuracy of the proposed estimation model in assessing breakage and impurity ratios, and evaluate the capability of MDSC-DeepLabv3+ in rapidly and accurately identifying the pixels of cane, broken cane, top, and leaf. Comprehensive experimental results show that the average relative errors of breakage and impurity ratio between predicted values and measured values are low. These findings have significant implications for the development of intelligent detection and cleaning system for sugarcane impurity.

2 Materials and methods

2.1 Raw sugarcane and impurity dataset construction

2.1.1 Detection device design

In order to provide a stable environment and meet the continuous image acquisition requirements that align with the raw sugarcane convey process in the sugar refinery, a dedicated platform for image acquisition of raw sugarcane and impurities was designed, as shown in Figure 1A. The platform mainly consists of portable energy storage, an acquisition room, a light source, an image acquisition module, a computer, and a motion assistance module.

The portable energy storage is used to supply power to the platform, especially in situations where electricity supply is limited. The interior of the image acquisition room, as depicted in Figure 1B, is covered with black matte paper to create a diffused lighting environment. Additionally, four magnetic base LED light bars are strategically placed around the room to ensure consistent illumination for the image acquisition module. The image acquisition module comprises an industrial camera and an industrial lens. The computer is connected to the image acquisition module via a USB 3.0 interface, which facilitates image storage and processing. The motion assistance module is composed of a conveyor, a cross beam guide rail, and a pair of vertical slider guide rails with self-locking function. The conveyor simulates the transmission of raw sugarcane before entering the pressing workshop. The vertical slider guide rails, equipped with scale markings, support and allow for adjustment of the cross beam guide rail where the camera is mounted. This feature enables easy adjustment of the camera's field of view and ensures the stability of the image acquisition module.

Table 1 shows the model parameters of the main components of the acquisition platform. The conveyor belt speed is determined based on sugar refinery practice and is measured in meters per second (m/s). The dimensions of the indoor acquisition room are set according to the requirements, with horizontal (H_{FOV}) and vertical (V_{FOV}) dimensions are set to the belt width of 450mm and indoor length of 600mm, respectively. The selected industrial camera has a horizontal (H_{CMOS}) and vertical (V_{CMOS}) size of the image sensor as 7.6×5.7mm, and the working distance (W_D) is set to 490mm considering the inner height of the acquisition room. The imaging principle of this acquisition platform is illustrated in Figure 2. Using the imaging principle and the dimensions of H_{CMOS}, V_{CMOS} and W_D, the field of view can be determined using Eq.(1).

$$f/W_D = V_{CMOS}/V_{FOV} = H_{CMOS}/H_{FOV} \quad (1)$$

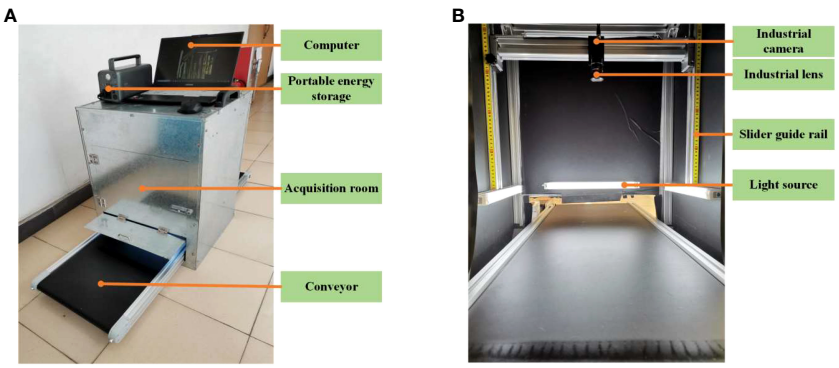


FIGURE 1
Machine vision acquisition platform. **(A)** Acquisition device structure. **(B)** Acquisition room.

TABLE 1 Main components of the acquisition platform.

Components	Parameters	Components	Parameters
Acquisition room	Indoor space 600mm×500mm×700 mm	Slider guide rail	SGR15N-500mm×2
Industry camera	MV-CA020-10UC with 89.1fps@1624×1240, image sensor size 7.6×5.7mm	Computer	AMD Ryzen7 5800H GeForce GTX 1650
Industry lens	MVL-MF0828M-8MP	Portable energy storage	72000mAh/3.2V
Light source	3600Lux×4	Conveyor	2000mm×450mm×100mm,1.5m/s, ≤20kg

As a result, the focal length is determined by $f = W_D \times (V_{CMOS}/V_{FOV}) = 490 \times (7.6/3450) = 8.27$mm, and the MVL-MF0828M-8MP industry lens is selected.

2.1.2 Image and mass data acquisition

The image and mass acquisition of raw sugarcane and impurities took place in the sugarcane unloading workshop of Junshi sugar refinery in Jijia Town, Leizhou City, Guangdong Province. The data collection period started from the middle of February to the end of the month in 2023, coinciding with the local sugarcane harvesting season. For this study, large-scale cultivated sugarcane variety "Yuetang 159" was selected. The raw sugarcane samples were randomly collected from different machine-harvested vehicles at various time intervals throughout the day using a loader. These samples were then manually placed on the conveyor belt of the acquisition platform for image collection. In total, 910 RGB 8-bit photos with jpg format and a resolution of 1624×1240 were captured. Each image contains four categories: cane, broken cane, top, and leaf, as shown in Figure 3. Following the image capturing process, 300 samples of raw sugarcane and impurities were randomly selected from the collected images. Each category of material in these samples was weighed using a calibrated electronic scale with a precision of 0.01g, and their masses were measured in grams (g).

2.1.3 Image labeling and dataset augmentation

The original dataset consists of 910 images containing cane, broken cane, top, leaf, and the background. These images were manually labeled and colored using the image annotation tool Labelme. The labeled regions of the five classes of objects were used to evaluate the training loss of intersection over union (IoU) between predicted bounding boxes and ground truth. The RGB values for cane, broken cane, top, and leaf were set to [128,0,0], [0,0,128], [0,128,0], and [128,128,0], respectively, while the background was set to [0,0,0]. To ensure model performance validation and testing, the dataset was randomly divided into training (546 images), validation (182 images), and test sets (182 images) with a ratio of 6:2:2.

In order to improve the generalization of the model, data augmentation techniques were applied to the training, validation, and test sets separately. Techniques such as random rotation, affine transformation, fogging, Gaussian noise, median filtering, and cutout were used to enhance the original images. After augmentation, the images were checked and corrected using Labelme to ensure accurate labeling of each class in every image. The annotated images were stored in the PASCAL VOC format and named Raw Sugarcane and Impurity (RSI). The label counting algorithm was used to calculate the number of labels in the RSI images, and the corresponding statistics are shown in Table 2. The dataset demonstrates a relatively balanced distribution of samples across each class. Examples of the original annotated images and augmented images can be observed in Figure 4.

2.2 Estimation model establishment

2.2.1 Surface density distribution analysis

In general, previous estimation models that are based on image pixels for assessing the mass of crops (such as wheat, corn, and soybean) often assume that the surface density (mass/pixel) of each crop category remains stable across different images (Chen et al., 2022; Jin et al., 2022; Liu et al., 2022). However, when it comes to broken cane and impurities, their surface density can vary significantly in different images. Therefore, before building the estimation model, it is essential to analyze the surface density distributions of cane, broken cane, top, and leaf separately. This analysis will help to account for the variation in surface density and ensure more accurate estimation for breakage and impurity ratios in raw sugarcane.

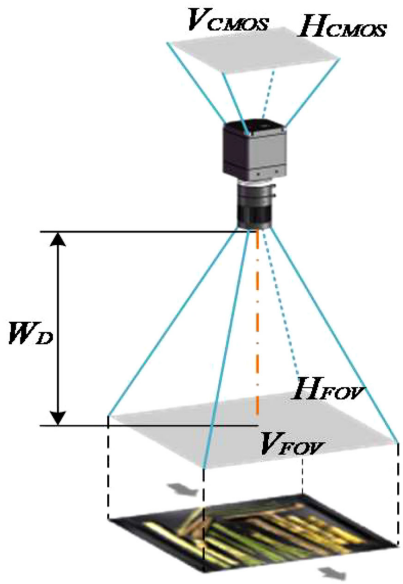


FIGURE 2
Imaging principle in this acquisition platform.

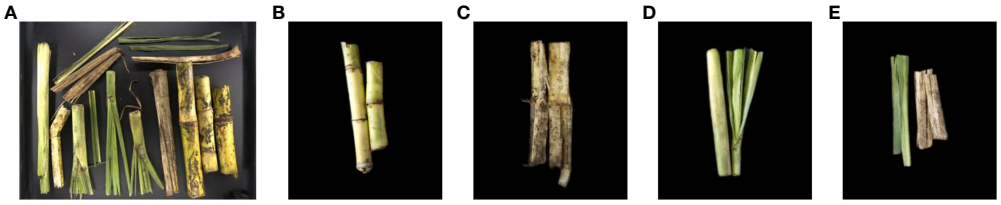


FIGURE 3
Acquisition materials and segmentation classes. **(A)** Original image, **(B)** Cane, **(C)** Broken cane, **(D)** Top, **(E)** Leaf.

The analysis of surface density distribution was conducted using 300 samples of mass data and the corresponding images for each category. The OpenCV threshold function was utilized to count the number of pixels in each category. Let P_C, P_B, P_T andP_L represent the number of pixels of cane, broken cane, top, and leaf in each image sample, respectively, and their corresponding masses are denoted as M_C, M_B, M_T and M_L, respectively. The spatial distribution of the surface density for raw sugarcane, including cane and broken cane, as well as the top and leaf, is presented in Figure 5. Based on the surface density distribution of raw sugarcane in Figure 5A, it can be observed that the surface density of cane fluctuates less and is more concentrated. The surface density of broken cane is approximately half of that of cane, and the data is scattered. Figures 5B, C illustrate that the surface density distribution of top and leaf is more scattered compared to broken cane.

To address the scattered surface density of broken cane, top, and leaf, a Gaussian distribution probability density function was used to fit the frequency histograms of surface density for each category. The mean surface density μ for each category was then obtained through the fitting process, and the results are demonstrated in Figure 6. It can be observed that all fitting coefficients R^2 are greater than 0.95, indicating high fitting accuracy.

The fitting results showed that the mean surface density of cane, broken cane, top, and leaf are μ_c = 1.52E-3, μ_B 7.4E-4, μ_T = 8.8E-4 and μ_L = 3E-5 with unit g/pix, respectively. Moreover, it is evident that the mean value of cane μ_c is approximately twice the mean value of broken cane μ_B and top surface density μ_T, and μ_c is more

TABLE 2 Statistic of Raw Sugarcane and Impurity (RSI) dataset.

Dataset	Training dataset	Validation dataset	Test dataset	Complete dataset
Images	5460	1820	1820	9100
Cane labels	16882	4151	3850	24883
Broken cane labels	13735	3310	3410	17045
Top labels	15903	4071	4390	24364
Leaf labels	17234	4015	3830	25079

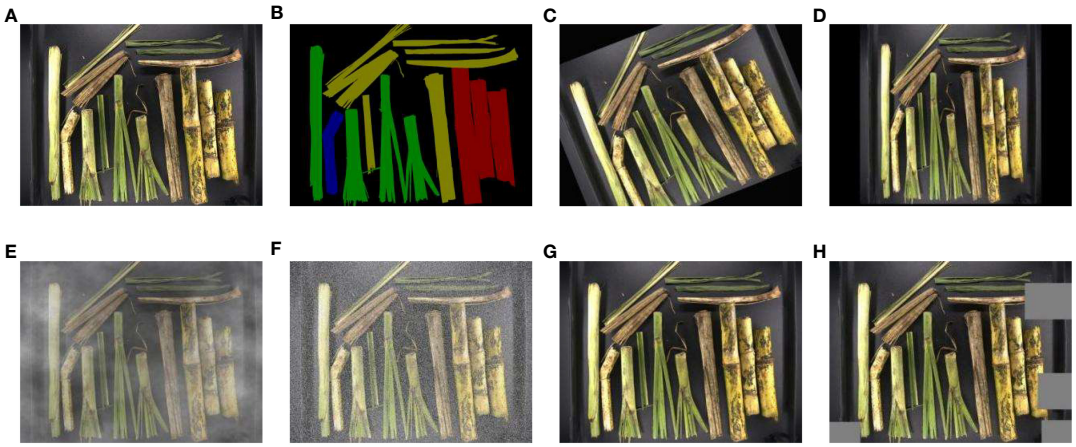


FIGURE 4
Augmented image samples and image label. **(A)** Original image, **(B)** Ground truth, **(C)** Random rotation, **(D)** Affine transformation, **(E)** Fogging, **(F)** Gaussian noise, **(G)** Median filtering, **(H)** Cutout.

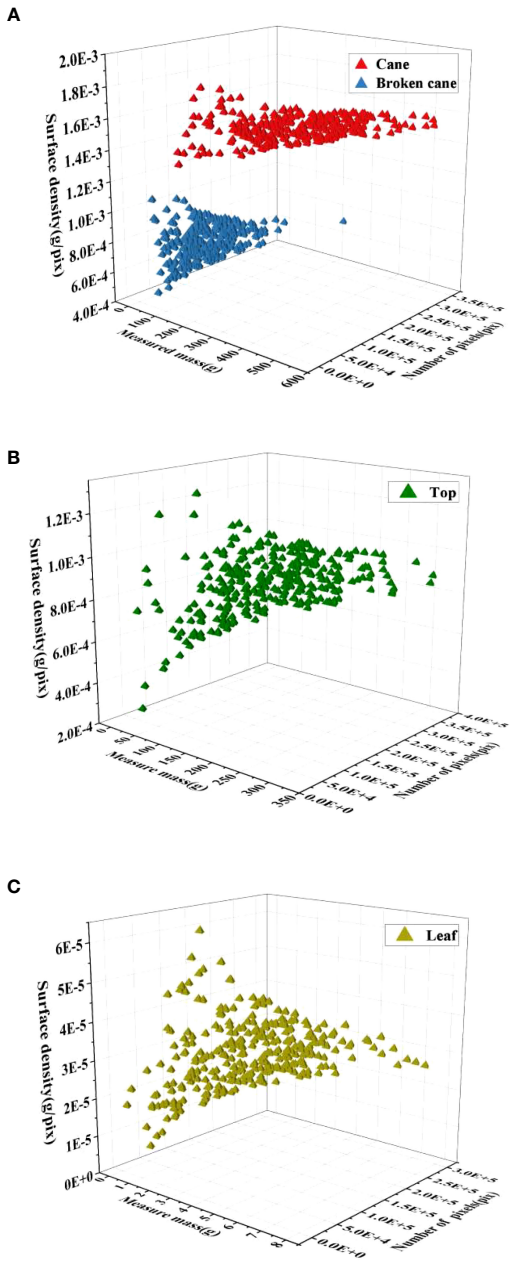


FIGURE 5
Spatial distribution of surface density for the 4 classes in RSI. **(A)** Raw sugarcane, **(B)** Top, **(C)** Leaf.

than fifty times *of* μ_L. The mass error of leaf has little effect on the overall mass error. Therefore, when establishing the estimation model, the accuracy of the estimated mass of cane should be ensured first, followed by broken cane, top, and finally leaf. This approach is consistent with the low deduction percentage setting (as low as 0.2%) employed by sugar refineries for leaf impurities.

2.2.2 Fitting and estimation model establishment

On the basis of the mean values of surface density given in Figure 6, the estimated mass of cane M'_C, broken cane M'_B, top M'_T, and leaf M'_L based on their pixels can be expressed as follows:

$$M'_C = \mu_C \times P_C = 1.52\text{E} - 3P_C \quad (1)$$

$$M'_B = \mu_B \times P_B = 7.4\text{E} - 4P_B \quad (2)$$

$$M'_T = \mu_T \times P_T = 8.8\text{E} - 4P_T \quad (3)$$

$$M'_L = \mu_L \times P_L = 3\text{E} - 5P_L \quad (4)$$

Furthermore, a linear regression of the estimated and measured mass was conducted to validate the accuracy of the mass estimation model defined by Eq.(1)-(4). Based on the distribution characteristics shown in Figure 6, a total of 285 mass data of cane, broken cane, top, and leaf within a 95% confidence interval were selected for fitting, and the fitting results were presented in Figure 7 and Table 3. It can be seen that the measured mass of the cane is highly correlated with the estimated mass with an R^2 value of 0.983. This indicates that the linear regression model is capable of explaining the numerical relationship between the measured mass and the estimated mass of the cane. The R^2 value for broken cane and top are 0.894 and 0.88, respectively, demonstrating the regression model's good fitting capability. The R^2 value for the leaf is 0.764 suggesting that the model can still adequately fit the relationship between the measured mass and the estimated mass. In addition, the results of ANOVA in Table 3 indicate that the significance F<0.01 between estimated cane, broken cane, top, and leaf and their measured values proves a high correlation.

Based on the mass of each category, the ratios of breakage (R_B) and impurity (R_I) *is* defined as:

$$\begin{aligned} R_B &= \frac{M_B}{M_C+M_B} \times 100\,\% \\ &= \frac{7.4E{-}4 \times P_E}{1.52E{-}3\times P_c+7.4E{-}4\times P_E} \times 100\,\% \end{aligned} \quad (5)$$

$$\begin{aligned} R_I &= \frac{M_T + M_L}{M_C+M_B+M_T+M_L} \times 100\,\% \\ &= \frac{8.8E{-}4\times P_T + 3E{-}5\times P_L}{1.52E{-}3\times P_c+7.4E{-}4 \times P_E+8.8E{-}4\times P_T + 3E{-}5\times P_L} \times 100\,\% \end{aligned} \quad (6)$$

Where M_C, M_B, M_T and M_L is the mass of cane, broken cane, top and leaf in an image sample. The estimated breakage and impurity ratios R'_B and R'_I can also be determined by replacing M_C, M_B, M_T and M_L in Eq.(5)-(6) with estimated mass M'_C, M'_B, M'_T and M'_L. Thereby Eq.(5)-(6) can be taken as the estimation model for breakage and impurity ratios.

2.3 Raw sugarcane and impurity segmentation model development

2.3.1 MDSC-DeepLabv3+ framework

In order to facilitate the M'_C, M'_B, M'_T, M'_L, R'_B and R'_I calculation, a segmentation model, MDSC-DeepLabv3+, was developed for the intelligent extraction of pixels of cane P_C, broken cane P_B, top P_T, and leaf P_L in each image sample. MDSC-DeepLabv3+ is an improvement upon the DeepLabv3+. The DeepLabv3+ comprises two modules: an encoder and a decoder (Chen et al., 2018). In the encoder, the Xception

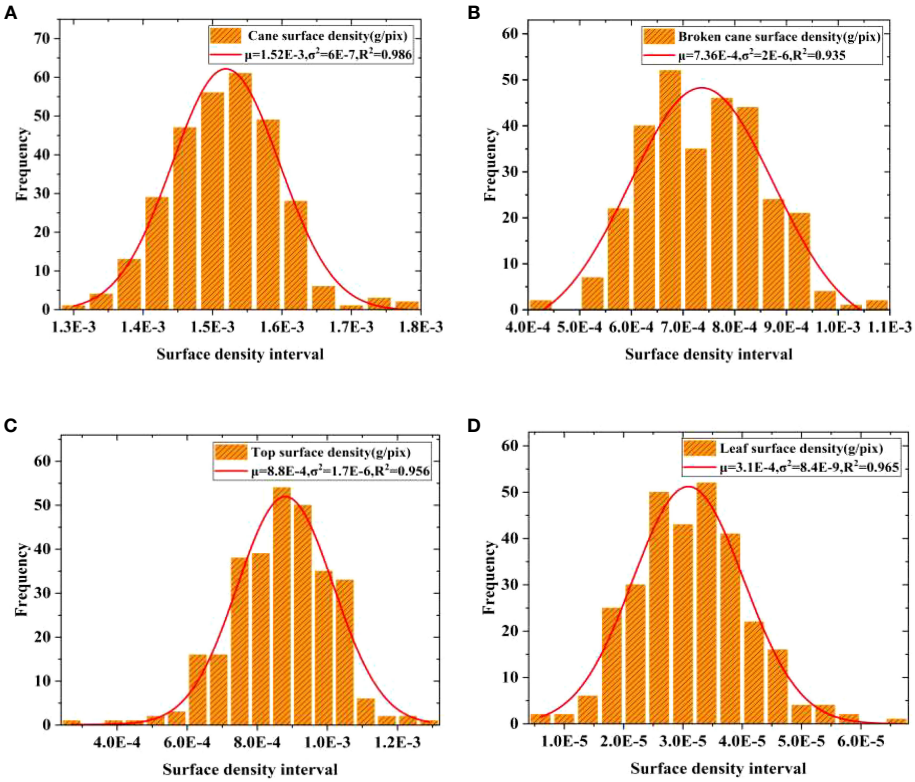


FIGURE 6
Gaussian distribution fitting of surface density. **(A)** Cane, **(B)** Broken cane, **(C)** Top, **(D)** Leaf.

backbone is used to extract input image features, resulting in two effective feature maps. One of the feature map undergoes processing through atrous spatial pyramid pooling named ASPP, and is then using a 1×1 standardization convolution for the fused features from ASPP. This produces high-level features that are subsequently fed into the decoder. The other feature map directly outputs to the decoder. The ASPP is composed of a 1×1 standardization convolution, three 3×3 depthwise separable convolutions named DSC with varying dilation rates (6, 12, and 18), and an average pooling layer. These convolutions generate feature maps at four different scales, which are stacked along the channel dimension.

In the decoder, the low-level features obtained from the Xception backbone first undergo 1×1 convolution to reduce the number of channels. Meanwhile, the high-level features from the encoder are bilinearly upsampled by a factor 4 to improve the image resolution. Afterwards, the 1×1 convoluted low-level features are fused with the upsampled high-level features, and a 3×3 DSC is utilized to extract information from the fused features, followed by another bilinear upsampling by a factor 4. Previous studies have demonstrated the effective use of DeepLabv3+ in agricultural fields, such as fruit picking, crop disease and pest, and field road scenes (Wu et al., 2021; Peng et al., 2023; Yu J. et al., 2023).

To enhance both the accuracy and deployability of the model, as well as reduce inference time, various improvements including improved MobileNetv2, ASPP_DS module and CA mechanism were introduced in this study. First, the atrous convolution was employed to optimize the MobileNetv2, and Xception was replaced by the improved MobileNetv2 in DeepLabv3+. In the MobileNetv2, dilated convolution was incorporated into the last two layers by increasing the kernel size, thus expanding the receptive field. This enhancement allows the network to better perceive surrounding information without significantly increasing computational complexity or compromising the resolution of the feature maps. Then, the dilation rates in the ASPP module were adjusted as 4, 8, and 12, and a strip pooling layer was added parallel to DSC to build a module named ASPP_DS. Module ASPP_DS can reduce the model parameters and establish long-range dependencies between regions distributed discretely, and focus on capturing local details. ASPP employs diverse padding and compact dilation strategies to extract receptive fields at various scales, effectively capturing information from both multi-scale contexts and small objects. Additionally, ASPP integrates a parallel strip pooling layer with elongated and narrow pooling kernels to grasp local contextual details in both horizontal and vertical spatial dimensions. This approach helps in reducing interference from unrelated regions in label prediction results. Finally, CA was appended to the output of MobileNetv2 and ASPP_DS separately, that allows the model to acquire weight information from the dimensions of feature channels and effectively leverage positional data. This incorporation enables the accurate capture of spatial relationships and contextual information of the target, thereby enhancing training efficiency. The enhanced version of DeepLabv3+ is denoted as MDSC-

10.3389/fpls.2023.1283230

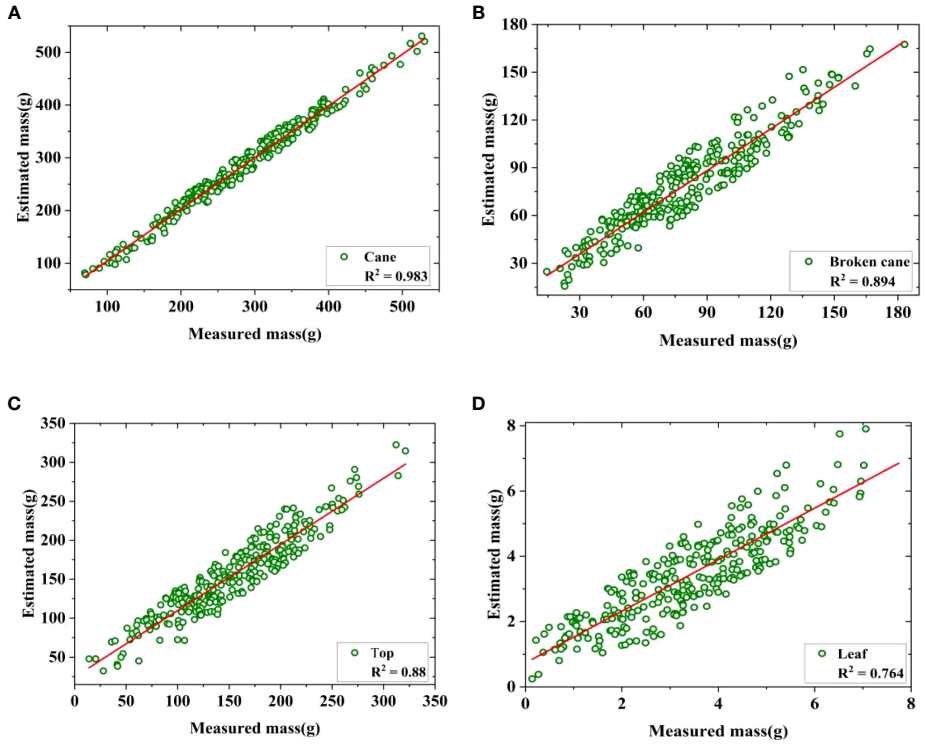


FIGURE 7
Regression of estimated and measured mass. **(A)** Cane, **(B)** Broken cane, **(C)** Top, **(D)** Leaf.

DeepLabv3+. The overall framework of MDSC-DeepLabv3+ is depicted in Figure 8.

2.3.2 Improved MobileNetv2

The basic structure unit of MobileNetv2 is the inverted residual block (IRB), which mainly consists of dimensionality expansion, feature extraction and dimensionality compress three main steps. The MobileNetv2 employs 3×3 depthwise convolution (Dwise) and 1×1 convolution to construct two IRBs with s= 1, s=2 (Sandler et al., 2018). In cases where the stride is equal to 1 and the shape of the input feature matrix matches that of the output feature matrix, a shortcut connection is employed, as shown in Figure 9. In addition, the dimensionality compression process in MobileNetv2 uses a linear activation function instead of the Relu activation function to reduce information loss caused by compression.

TABLE 3 Analysis of Variance (ANOVA) of estimated and measured mass. .

Category		DF	Square sums	Mean square	F	Significance F
Cane	Regression analysis	1	2340192.15697	2340192.15697	16820.25846	4.23041E-254
	Residual	283	39373.61497	139.12938		
	Total	284	2379565.77194			
Broken cane	Regression analysis	1	225202.9665	225202.9665	2390.43448	4.97988E-140
	Residual	283	26661.44583	94.21006		
	Total	284	251864.41233			
Top	Regression analysis	1	656015.70993	656015.70993	2055.21929	8.58987E-132
	Residual	283	90332.18347	319.19499		
	Total	284	746347.8934			
Leaf	Regression analysis	1	431.20971	431.20971	915.53104	1.07792E-90
	Residual	283	133.29133	0.47099		
	Total	284	564.50104			

10.3389/fpls.2023.1283230

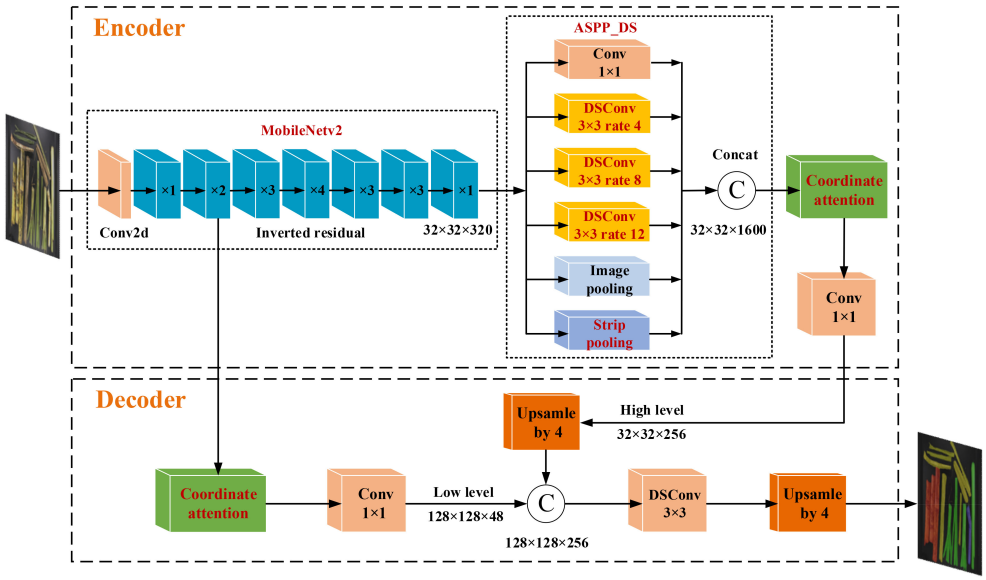


FIGURE 8
Framework of MDSC-DeepLabv3+.

To reduce computing costs and memory usage, this study utilizes the first 8 layers of the MobileNetv2 model. This choice is made because starting from the 9th layer, the number of output channel increases to 1280, leading to higher computing resource consumption. To minimize the loss of down-sampling information while increasing receptive field, the stride of the 7th layer is modified to 1 (Meng et al., 2020).

Furthermore, dilated convolutions with a factor not exceeding 1 are utilized to replace conventional convolutions. According to research by Wang et al. (2018), sparse concatenation of dilated convolution may introduce grid effects, hindering the lower layers of the network from fully leveraging features from the initial layer and causing the loss of fine-grained details. Therefore, dilation rates of 2 and 5 are applied in the 7th and 8th layer respectively, while the remaining layers maintain a dilation rate of 1, aiming to expand the receptive field and preserve edge detail information. The structure and hyperparameter of the improved MobileNetv2 are displayed in Table 4, in which *t* is the expansion factor, *c* is the output channel, n is the number of repetitions of bottleneck, *s* is the first module strid, and *r* is dilation rate. When dilation rate of 1 results in atrous convolution being equivalent to a regular convolution. This design achieves a balance between computational resource consumption and network performance requirements.

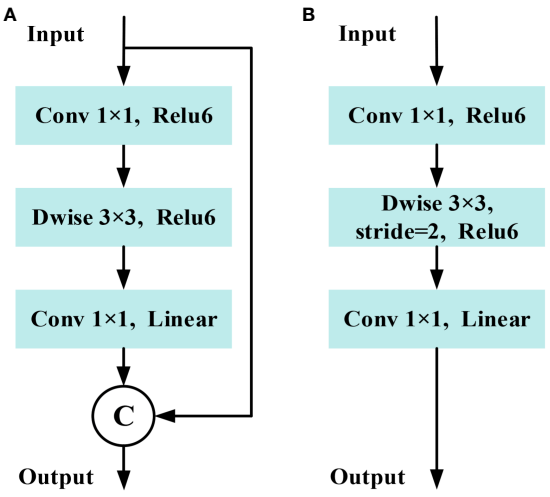


FIGURE 9
Structure of inverted residual block in MobileNetv2. **(A)** Stride=1 block. **(B)** Stride=2 block.

2.3.3 Strip pooling

To better handle the segmentation of broken cane and top with irregular and complex shapes, a lightweight strip pooling layer was added in parallel to DSC in the ASPP. This allows for more efficient acquisition of information from a large receptive field, facilitating the collection of remote contextual information from different spatial dimensions by ASPP. Strip pooling utilizes a pooling kernel (rectangular area) that performs pooling operations along the horizontal and vertical dimensions. The structure of strip pooling (Hou et al., 2020) is shown in Figure 10, where X ϵ $R^{C\times H\times W}$ is the input tensor, C denotes the number of channels, H denotes the height, and W denotes the width. First, the input X is pooled horizontally and vertically to obtain y^h ϵ $R^{C\times H\times 1}$ and y^v ϵ $R^{C\times 1\times W}$, respectively. Then, the feature maps are expanded to the same resolution C×H×W as the input X using a 1D convolution with a kernel size of 3×3 to obtain the expanded y^h, y^v. Next, the expanded feature maps are fused to obtain a final representation.

$$y_{c,i,j} = y^h_{c,i} + y^v_{c,j}, 1 \le c \le C, 1 \le i \le H, 1 \le j \le W$$

Finally, after a 1 × 1 standard convolution and a sigmoid layer, the final output Z of strip pooling is obtained by multiplying the corresponding elements with the original input.

$$Z = Scale(X, \sigma(f(y)))$$

where Scale (-, -) is the element-level multiplication, σ is the sigmoid function, and f is the 1×1 convolution, y is feature fusion results.

The element of specified location in the output tensor (i, j),$1 \le i \le H$,$1 \le j \le W$ corresponds to the result of strip pooling of the horizontal and the vertical pooling window in the input tensor. By repeatedly applying the aggregation process using long and narrow pooling kernels, the ASPP_DS module can efficiently capture information from a wide receptive field throughout the entire scene. Due to the design of the elongated and narrow shape of the pooling kernel, it not only establishes remote dependency relationships between regions distributed discretely but also focuses on capturing local detailed features.

2.3.4 Coordinate attention

Inspired by the prominence of the region-of-interest search in the human visual system, attention mechanisms aim to simulate this process by dynamically adjusting the weights based on the input image features. Attention mechanisms can be categorized into various types, such as channel attention (e.g. SE), hybrid attention (e.g. CBAM), temporal attention (e.g. GLTR), branch attention (e.g. SKNet), and position attention mechanisms (e.g. CA). These attention mechanisms have been widely applied in fields such as object detection (Yu J. et al., 2023) and image segmentation (Zhu et al., 2023).

TABLE 4 Hyperparameters of MobileNetv2.

Input size	Operator	t	c	n	s	r
512×512×3	conv2d	–	32	1	2	1
256×256×32	bottleneck	1	16	1	1	1
256×256×16	bottleneck	6	24	2	2	1
128×128×24	bottleneck	6	32	3	2	1
64×64×32	bottleneck	6	64	4	2	1
32×32×64	bottleneck	6	96	3	1	1
32×32×96	bottleneck	6	160	3	1	2
32×32×160	bottleneck	6	320	1	1	5

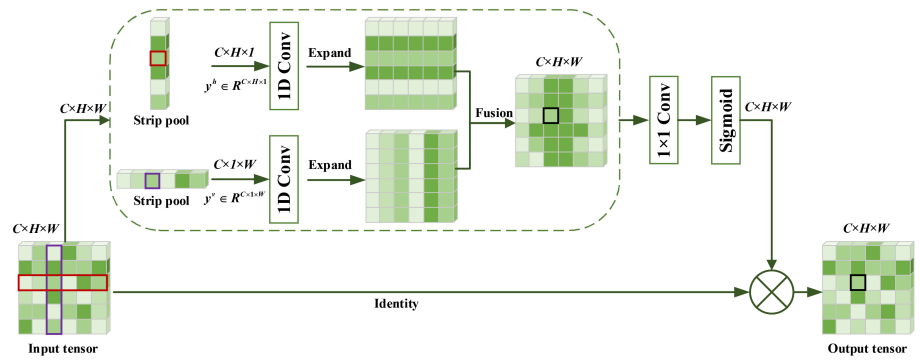


FIGURE 10
Structure of strip pooling.

The CA not only models channel relationships but also utilizes positional information to capture long-range dependencies (Hou et al., 2021). Therefore, CA was selected in the MDSC-DeepLabv3+ to highlight the regions of interest. The CA consists of coordinate information embedding (CIE) and coordinate attention generation (CAG) two main operation, as shown in Figure 11. CIE introduces two global average pooling to encode each channel along the horizontal and vertical coordinate on the input feature map, respectively, hence aggregates features along the two spatial directions. These two pairs of global average pooling operation enable CA to capture long-range dependencies along one spatial direction and preserve precise positional information along other one, which allows the network to more precisely locate the objects of interest. CAG first conducts concatenation (Concat) and Conv2d for the feature maps obtained from CIE followed by batch normalization and non-linear activation operation. Then, the intermediate feature map is split into two separate tensors along the spatial dimension. Next, 1×1 Conv2d and sigmoid activation are utilized to separately transform the output tensors to tensors with the same channel number as the input feature maps. Finally, the output tensors are then expanded into elements and used as attention weights. The final output of CA is the element-wise multiplication of original input of CIE and the attention weights.

Introduction of CA before low feature processing and after the features fusion of ASPP_DS is beneficial in fully utilizing positional information. This allows the model to accurately capture the spatial relationships and contextual information of the target, thus improving the accuracy of sugarcane and impurity phenotype segmentation in denser images.

3 Experiments and results

3.1 Analyzing of estimation model

The effectiveness of estimation model for breakage and impurity ratios defined in Section 2.2.2 was validated by fitting estimated and measured value. First, the measured mass of cane, broken cane, top, and cane leaf M_C, M_B, M_T and M_L, along with the number of pixels for each category manually labeled in the selected 285 images (95% confidence interval of samples) were obtained. Then, estimated masses of M'_C, M'_B, M'_T and M'_L for the four categories were determined based on the mean surface density μ_C, μ_B, μ_T and μ_L according to Eq.(1)-(4). Next, the measured and estimated ratios of breakage and impurity were obtained according to Eq.(5)-(6) based on the measured and estimated masses. Finally, the measured breakage and impurity ratios were linearly fitted with the estimated breakage and impurity ratios, and the fitting results are shown in Figure 12 and Table 5, respectively.

It can be observed that the fitting R^2 values are as high as 0.976 and 0.968, respectively. In addition, the results of the ANOVA presented in Table 5 indicate a high correlation between the estimated breakage and impurity ratios and their measured values, with a significance level of $F<0.01$. Therefore, it is feasible to utilize the fitted surface density to estimate mass for each category and furthermore predict the breakage and impurity ratios for raw sugarcane.

3.2 Analyzing of segmentation model

3.2.1 Training environment and evaluation metrics

The semantic segmentation categories considered in this study are background, cane, broken cane, top, and leaf. In the process of sugarcane harvesting, raw sugarcane is primarily composed of cane, with cane tops and leaves present as impurities to a lesser extent. Broken cane represents the category with the lowest representation, leading to an extreme class imbalance. Consequently, this often leads to imbalanced positive and negative samples, along with varying sample difficulties. Therefore, this study utilizes the Focal Loss function as the primary loss function to address the imbalance between easy and difficult samples, facilitating better parameter optimization during the backpropagation process (Lin et al., 2017). In addition, the model incorporates the multi-class Dice Loss as an auxiliary loss function to enhance segmentation accuracy and address class imbalance scenarios (Milletari et al., 2016). The combination of Focal Loss and multi-class Dice Loss as the loss function enhances the model's predictive capability. The Focal loss for multi-objective segmentation is defined as.

$$L_F = -\alpha_t(1-p_t)^{\gamma}\log(p_t)$$

Where p_t is the confidence value of the sample category prediction. γ is an adjustable parameter, and the default is 2.

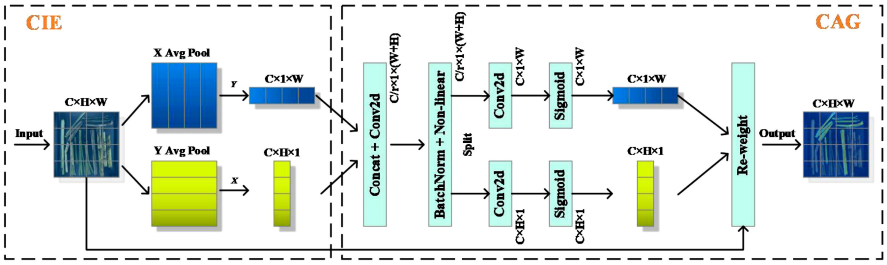


FIGURE 11
Structure of coordinate attention.

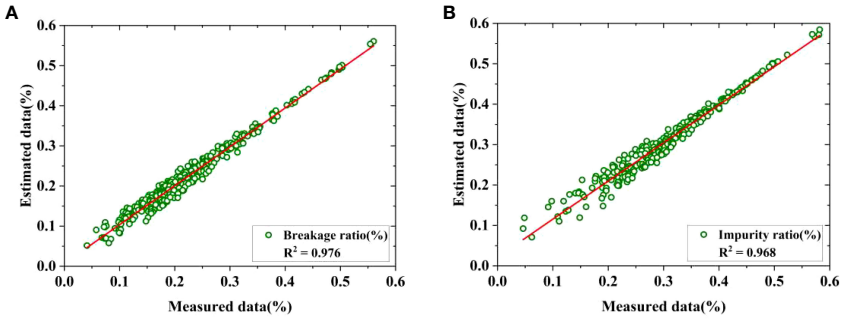


FIGURE 12
Fitting of estimated and measured ratio. **(A)** Breakage ratio, **(B)** Impurity ratio.

The Dice loss for multi-objective segmentation is defined as.

$$L_D = 1 - \sum_{j=1}^{c} \frac{2W_j \sum_{i=1}^{N} gt(j,i)\log(p_{i,j})}{\sum_{i=1}^{N}\left(gt(j,i)^2 + \log(p_{i,j})^2\right)}$$

Where, N is the number of samples, c is the target class, and $p_{i,j}$ is the softmax output of class j target class; $gt(j,i)$ is the ground-truth label of class j target, and W_j is the weight of the objective of class j, $W_j = 1/j$.

The experiments were conducted on a server in the lab with the configuration shown in Table 6. The MDSC-DeepLabv3+ used the Adam optimizer to compute the gradient of the loss function in each epoch to perform parameter updates. The initial learning rate was set to E-4. The batch size was set to 6. The training process consists of 100 epochs. In each epoch, the image dataset was randomly shuffled and fed into the model to ensure a different order of dataset used in different epochs. This technique enhances the convergence speed of the model and improves the prediction results on the test set.

In order to comprehensively evaluate the performance of the proposed and comparative semantic segmentation models, three aspects of each model, namely accuracy, deployability, and efficiency, are comprehensively evaluated. The commonly used mIoU and mPA were utilized as accuracy evaluation metrics. And the model deployability was evaluated using model parameter quantity (Param) and model computation volume floating point operations (FLOPs). Efficiency was evaluated using inference time for each image. The metrics of IoU, mIoU and mPA which is represented by the following Eq. (7)-(9), respectively.

$$IoU_i = \frac{P_{ii}}{\sum_{i=0}^{c-1} P_{ij} + \sum_{j=0}^{c-1} P_{ji} - P_{ii}} \times 100\,\% \tag{7}$$

$$mIoU = \frac{1}{c}\sum_{i=0}^{c-1} IoU_i \tag{8}$$

$$mPA = \frac{1}{c}\sum_{i=0}^{c-1} \frac{p_{ii}}{\sum_{j=0}^{c-1} p_{ij}} \tag{9}$$

Where c denotes the number of categories, so c=4 (cane, broken cane, top and leaf), P_{ij} or P_{ji} denotes the number of category prediction that is incorrect, while P_{ii} denotes the number of correct predictions made by categories.

3.2.2 Model training

The size of the input image is a crucial factor affecting the model's performance. Increasing the image size enhances accuracy by preserving semantic information for small targets and preventing information loss caused by low-resolution feature maps. However, excessively large image sizes can lead to reduced detection accuracy due to the limited receptive field imposed by the fixed network structure. This, in turn, diminishes the network's ability to accurately predict targets of various scales (Lin et al., 2022). In practical applications, there is a trade-off between accuracy and speed that requires careful consideration. For this study, the input image was resized to three different dimensions: 256×256, 512×512, and 768×768. The proposed MDSC-DeepLabv3 + model was trained accordingly, and the results obtained are presented in Table 7. It can be observed that reducing the input

TABLE 5 ANOVA of breakage and impurity ratios.

Ratio		DF	Square sums	Mean square	F	Significance F
Breakage ratio	Regression analysis	1	2.58018	2.58018	11405.03085	1.05518E-230
	Residual	283	0.06402	2.26232E-4		
	Total	284	2.64421			
Impurity ratio	Regression analysis	1	2.41267	2.41267	8470.24579	6.21725E-213
	Residual	283	0.08061	2.84841E-4		
	Total	284	2.49328			

TABLE 6 Experimental environment.

Parameter	Configuration	Parameter	Configuration
Operating system	Ubuntu 18.04	Operating environment	CUDA 11.2
Deep learning framework	PyTorch 1.8	CPU	Intel(R) Xeon(R) Silver 4214 CPU @2.20GHz
Programming Language	Python 3.7	GPU	NVIDIA GeForce RTX 3080 12G @1260-1710MHz

image size to 512×512 achieves an optimal balance between speed and accuracy.

The segmentation results of models using different loss functions are displayed in Figure 13. The MDSC-DeepLabv3+ using only the Dice loss function exhibits the highest fluctuations in mPA and mIoU, leading to inferior segmentation results. Similarly, the MDSC-DeepLabv3+ using only Focal Loss demonstrates notable fluctuations during the early stages of the validation process, with slow growth in mPA and mIoU values in later stages. In contrast, the MDSC-DeepLabv3+ which combines Focal Loss and multi-class Dice Loss exhibits lesser sawtooth fluctuations during the increase in mPA and mIoU values, ultimately reaching their peak during the validation process. Consequently, the integration of Focal Loss and multi-class Dice Loss yields optimal outcomes in the segmentation of raw sugarcane and impurities.

3.2.3 Ablation experiment

To verify the effectiveness of the three improvements, including improved MobileNetv2, ASPP_DS and CA presented in Section 2.3, the following 7 models were constructed according to the control variable method, with a downsampling factor of 8.

TABLE 7 mPA and inference time obtained with different input image sizes.

Resize of image/pixels	mPA/%	Inference time/ms
256×256	94.68	10.69
512×512	97.55	13.85
768×768	97.07	24.19

1. DeepLabv3+_base: MobileNetv2 replaced the backbone Xception in DeepLabv3+.
2. M-DeepLabv3+: MobileNetv2 in DeepLabv3+_base was enhanced with atrous convolution operation.
3. MDS-DeepLabv3+: ASPP_DS replaced ASPP module in M-DeepLabv3+.
4. MC1-DeepLabv3+: CA was applied independently before 1×1 Conv of low-level features by the decoder in M-DeepLabv3+.
5. MC2-DeepLabv3+: CA was applied independently after the fusion of ASPP in M-DeepLabv3+.
6. MC-DeepLabv3+: CA was added separately before 1×1 Conv the low-level features and after the fusion of ASPP features in M-DeepLabv3+.
7. MDSC-DeepLabv3+: CA was added separately before processing the low-level features and after the fusion of ASPP_DS features in MDS-DeepLabv3+.

Table 8 presents the results of the ablation experiment for the seven aforementioned models. It can be observed that the MDSC-DeepLabv3+ outperforms the baseline DeepLabv3+_base, with an improvement of 1.25 in mPA and 1.8 in mIoU. Additionally, it achieves a reduction of 16.42% in Params and 31.46% in FLOPs, however, the inference time per image has slightly increased from 13.48ms to 13.85ms. These results demonstrate that the MDSC-DeepLabv3+ surpasses the DeepLabv3+_base in terms of segmentation accuracy and deployability metrics, while still maintaining comparable efficiency. Furthermore, it can be seen that the MDSC-DeepLabv3+ achieves the highest segmentation accuracy (mPA and mIoU) compared to other models, while exhibiting minimal differences in terms of deployability (Params, FLOPs) and efficiency (inference time) metrics.

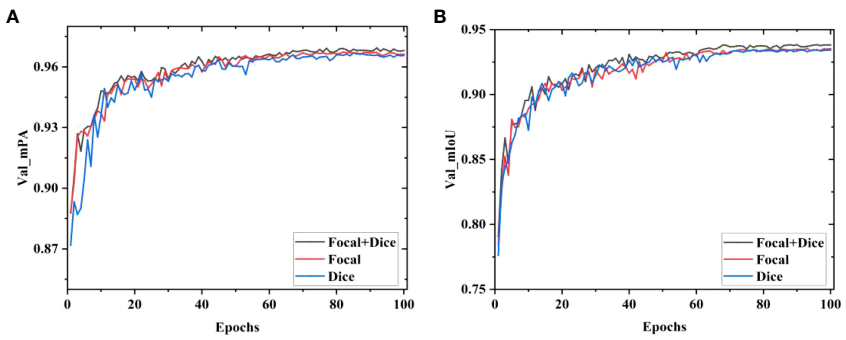


FIGURE 13
Results of mPA and mIoU with different loss functions. **(A)** Valid mPA, **(B)** Valid mIoU.

In order to visually demonstrate the improvement of the models, Grad-CAM (Selvaraju et al., 2020) was used to visualize the channels of the feature maps of DeepLabv3+ and MDSC-DeepLabv3+. The visualization segmentation instances of top were illustrated in Figure 14. In group (a), the two feature maps are extracted by the Xception in DeepLabv3+ and the enhanced MobileNetv2 in MDSC-DeepLabv3+, respectively. In group (b), two feature maps are the output of ASPP in DeepLabv3+ and ASPP_DS in MDSC-DeepLabv3+, respectively. In group (c), the two feature maps are the output of DeepLabv3+ and MDSC-DeepLabv3 +, respectively.

In Figure 14A, it can be observed that Xception in DeepLabv3+ achieves clearer pixel segmentation than that obtained by MobileNetv2 in MDSC-DeepLabv3+. The reason is that MobileNetv2 is a lightweight and shallow model compared to Xception, and its depthwise convolution can lead to information loss and limit the number of channels, thereby resulting in a lower-level feature map with fewer information. However, the two heat maps in group (b) indicate that there is pixels misfocus at the top-right corner in the first line of the feature map extracted by ASPP, while ASPP_DS results in more complete pixel segmentation, enhances preservation of details, and eliminates the top-right misfocus. The heat map illustrates that the introduced strip pooling in ASPP_DS rectifies the shortage of MobileNetv2, and the dense and compact dilation rates (4, 8, 12) improve its capability of focusing on capturing local detailed features. Heat map of final outputs of MDSC-DeepLabv3+ and DeepLabv3+ given in Figure 14C demonstrates that the CA in MDSC-DeepLabv3+ further enhances the color intensity in heat map, indicating that the inclusion of CA allows the model to focus more on the features of the categories, thereby enhancing its distinguishability of cane, broken cane, top and leaf.

3.2.4 Comparative experiment

To further validate the superiority of the proposed model MDSC-DeepLabv3+, comparative experiments were conducted using the RSI dataset under the same experimental conditions. The compared models include UNet, PSPNet, SegFormer-B0, and the baseline DeepLabv3+. Previous research results have shown that UNet (Ronneberger et al., 2015) and PSPNet (Zhao et al., 2017) perform well in terms of accuracy in segmentation tasks with challenges like cell tracking ISBI and Cityscapes. SegFormer-B0 is

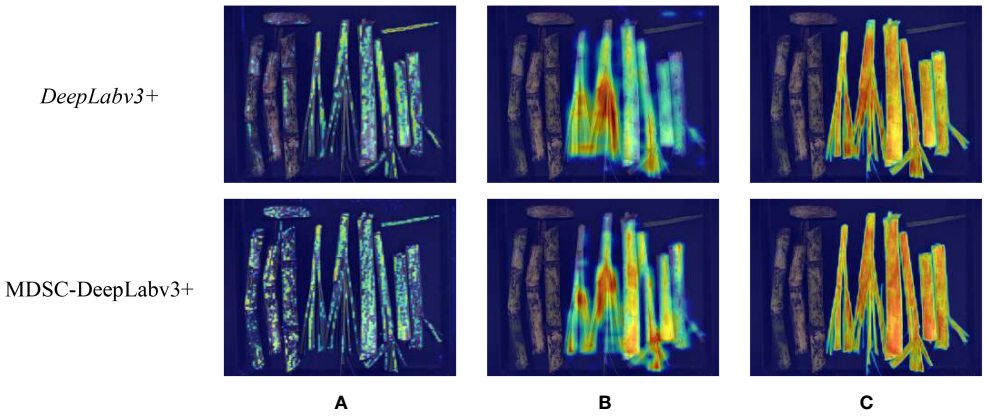


FIGURE 14
(A) Backbone output. **(B)** Encode output. **(C)** Decode output.

TABLE 8 Results of ablation experiment.

Number	ASPP_DS	Coordinate Attention		mPA/ %	mIoU/ %	Param/ M	FLOPs/ G	Inference time/ ms
		Before decoder	After ASPP (_DS)					
(1)				96.3	93.04	4.81	69.29	13.48
(2)				96.67	93.36	3.35	45.49	12.13
(3)	√			97.16	94.37	3.36	45.41	12.76
(4)		√		96.88	93.66	3.55	46.83	13.51
(5)			√	97.05	94.48	3.63	46.88	13.56
(6)		√	√	97.22	94.57	3.68	46.88	13.67
(7)	√	√	√	97.55	94.84	4.02	47.49	13.85

a lightweight model that combines transformers with a lightweight multilayer perceptron decoder (Xie et al., 2021). The comparative results are given in Table 9.

It can be seen that the accuracy of MDSC-DeepLabv3+ surpasses that of the aforementioned four models with significant improvements. Specifically, the mIoU of MDSC-DeepLabv3+ is higher by 0.81, 5.22, 12.47, and 0.28 compared to UNet, PSPNet, SegFormer-B0, and DeepLabv3+, respectively. Moreover, the mPA of MDSC-DeepLabv3 + reaches an impressive 97.55%, which outperforms UNet, PSPNet, SegFormer-B0, and DeepLabv3+ by 0.69, 2.7, 7.76, and 0.34, respectively. These remarkable improvements can be attributed to the adoption of the advanced DeepLabv3+ as the basic model, coupled with the enhancements introduced through strip pooling and CA. Strip pooling plays a crucial role in collecting remote contextual information from different spatial dimensions and addressing the issue of information loss resulting from the atrous convolution operation in DeepLabv3. On the other hand, CA efficiently utilizes positional information, enabling accurate capturing of the spatial relationships and contextual information of the detected cane, broken cane, top, and leaf.

In terms of deployability, MDSC-DeepLabv3+ demonstrates remarkable reductions in Params and FLOPs when compared to UNet, PSPNet, and DeepLabv3+. Specifically, it reduces Params by 83.65%, 91.29%, and 90.35%, and FLOPs by 89.49%, 59.9%, and 66.37% compared to UNet, PSPNet, and DeepLabv3+ respectively. This significant reduction in model size and computational complexity makes MDSC-DeepLabv3+ highly efficient and resource-friendly. Moreover, MDSC-DeepLabv3+ achieves impressive segmentation efficiency, with a recognition speed of only 13.85ms per image. This inference time per image is far less than the above three models, with reductions of 48.97%, 10.18%, and 43.31%, respectively. This indicates that MDSC-DeepLabv3+ is able to perform fast and accurate segmentation, making it highly suitable for real-time applications. Although SegFormer-B0 may have some advantages in terms of deployability, its accuracy is much lower compared to MDSC-DeepLabv3+ (89.79% vs. 97.55%). The reason for this superior performance is the utilization of the improved lightweight MobileNetv2, which replaces Xception in DeepLabv3+, leading to an efficient and accurate model overall. In summary, the proposed MDSC-DeepLabv3+ outperforms the compared four models in the task of segmenting sugarcane and impurities, offering a winning combination of high segmentation accuracy, deployability, and recognition speed.

Instances of the results obtained using the aforementioned segmentation models are illustrated in Figure 15. In which, red [128,0,0] represents cane, blue [0,0,128] represents broken cane, green [0,128,0] represents top, yellow [128,128,0] represents leaf, and black [0,0,0] represents the background. From the visualization of test results, it is evident that all five models perform well in most cases. However, the segmentation obtained by MDSC-DeepLabv3+ stands out as more complete, with clearer preservation of details in general. Upon closer observation, it can be seen that UNet, PSPNet, and SegFormer-B0 misclassify their categories, for instance, misclassifying broken cane as leaf, and vice versa. This indicates inaccuracies in pixel differentiation for these models. Additionally, the compared four models result in fuzzy segmentation and ambiguous boundaries between objects. On the other hand, the proposed MDSC-DeepLabv3+ demonstrates superior performance in addressing the issue of detail adhesion. This can be observed in the instances marked out in the line of MDSC-DeepLabv3+ where the model is capable of better distinguishing object boundaries and preserving fine details.

3.3 Analyzing of comprehensive experiment

The breakage and impurity ratios of raw sugarcane were estimated using the estimation model presented in Section 2.2 and the MDSC-DeepLabv3+ segmentation model presented in Section 2.3. These estimated values were then compared with the measured breakage and impurity ratios obtained through manual weighing to assess the effectiveness of the proposed method.

First, a subset of 25% (70) of the images was randomly selected from the mass dataset with 300 samples. The MDSC-DeepLabv3+ model was applied to semantically segment the selected 70 images and determine the number of cane, broken cane, top, and leaf pixels for each image. Then, corresponding masses were estimated using Eq.(1)-(4), based on the mean values of the surface density for each category obtained through normal fitting. The ratios of breakage and impurity were calculated according to the estimation model defined in Eq.(5)-(6) based on the estimated masses. Finally, the measured breakage and impurity ratios were determined using the measured mass and the relative errors between the estimated and measured results were calculated. Tables 10. 11 document and analyze the relative errors in the breakage ratio and impurity ratio for each sample, as well as the

TABLE 9 Test results of different recognition models.

Segmentation models	IoU/%					mIoU/%	mPA%	Param/M	FLOPs/G	Inference time/ms
	Background	Cane	Broken cane	Top	Leaf					
UNet	98.13	94.18	91.01	93.11	93.73	94.03	96.86	24.89	451.77	27.14
PSPNet	95.45	90.38	87.89	86.48	87.89	89.62	94.85	46.71	118.43	15.42
SegFormer-B0	95.6	82.98	72.38	81.12	79.79	82.37	89.79	3.72	13.56	16.78
DeepLabv3+	97.78	95.18	91.83	93.38	94.62	94.56	97.21	42.19	141.22	24.43
MDSC-DeepLabv3+	97.94	95.13	91.85	94.27	95.03	94.84	97.55	4.07	47.49	13.85

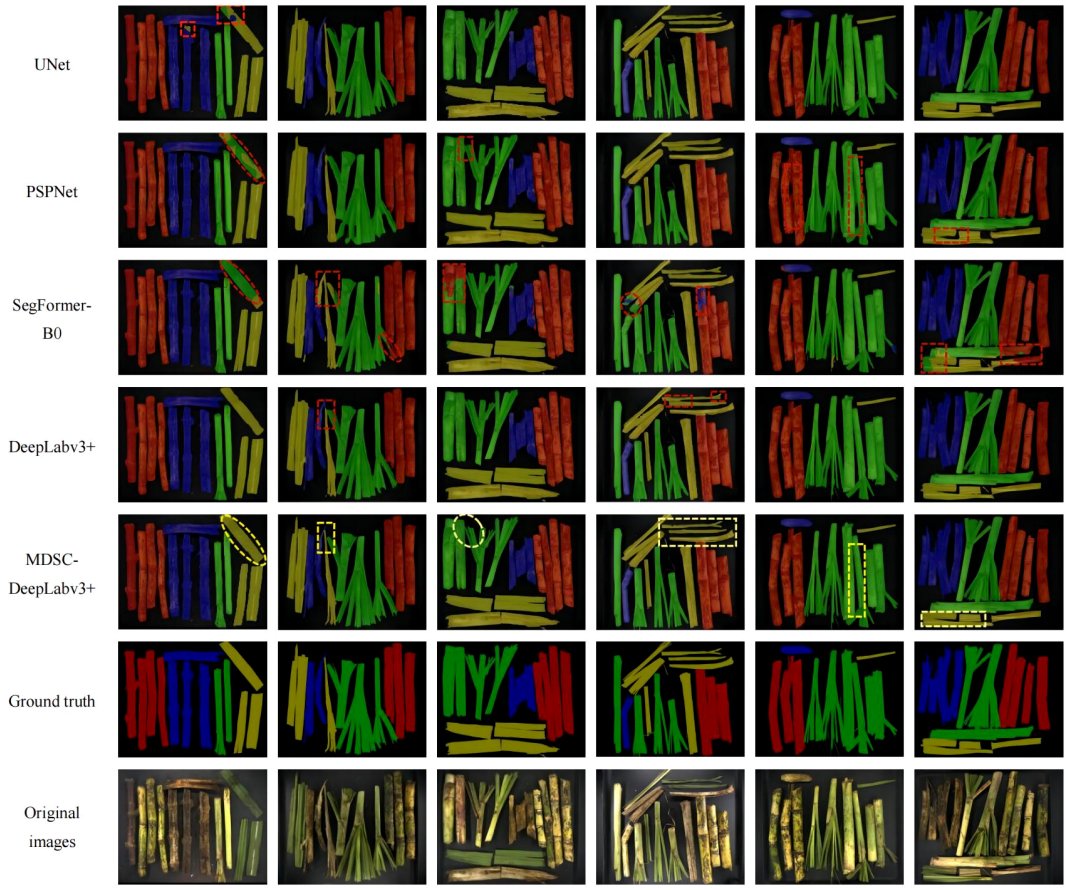


FIGURE 15
Test results of each detection model.

average relative error of the overall samples. The average relative errors were found to be 11.3% and 6.5% for breakage and impurity ratios, respectively. These results indicated that the proposed method exhibits strong reliability.

Additionally, the visualization of measured and estimated ratios of the 70 samples is depicted in Figure 16. This aids in the intuitive observation and analysis of the relationship and differences between predicted and manual measured results. It can be observed that the results obtained using the proposed method exhibit only slight deviations compared to the results obtained through manual weighing measurements, and the fluctuations are minimal. This suggests that the estimated breakage and impurity ratios can maintain their stability. Consequently, the proposed method based on estimation model and MDSC-DeepLabv3+ offers an efficient, accurate, and intelligent means of quantitatively estimating the breakage and impurity ratios of raw sugarcane.

4 Conclusions

In practice, objective, efficient, accurate, and intelligent detection of breakage and impurity ratios is an urgent requirement in the sugar refinery. Therefore, this study developed a novel approach combining the estimation model and MDSC-DeepLabv3+ segmentation network to tackle this problem. First, a machine vision-based acquisition platform was designed, and custom image and mass datasets of raw sugarcane and impurities were constructed. Then, estimation model was built to assess the ratios of breakage and impurity, considering the variation of surface density for the four categories of objects. Finally, the MDSC-DeepLabv3+ segmentation network dedicated to the detection of cane, broken cane, top, and leaf was developed. It effectively incorporated improved MobileNetv2, ASPP_DS, and CA based on DeepLabv3+ to enhance segmentation accuracy, reduce parameters and inference time. The analysis of the experimental results leads to the following conclusions:

1. The breakage and impurity ratios obtained through estimation model based on normal fitted surface density exhibit high accuracy, with corresponding R^2 of 0.976 and 0.968, respectively.
2. The proposed MDSC-DeepLabv3+ achieved superiority considering segmentation accuracy, deployability, and efficiency simultaneously. The mPA and mIoU achieved by MDSC-DeepLabv3+ were as high as 97.55% and 94.84%, respectively, surpassing the baseline DeepLabv3+ by 0.34

10.3389/fpls.2023.1283230

TABLE 10 Breakage ratios of 70 samples.

Sample number	Breakage ratio/%			Sample number	Breakage ratio/%		
	Measured	Estimated	Relative errors		Measured	Estimated	Relative errors
1	0.393	0.433	0.103	36	0.163	0.146	0.103
2	0.244	0.215	0.119	37	0.067	0.077	0.148
3	0.328	0.319	0.027	38	0.075	0.077	0.017
4	0.122	0.096	0.218	39	0.117	0.127	0.082
5	0.259	0.272	0.049	40	0.253	0.242	0.047
6	0.486	0.562	0.156	41	0.381	0.418	0.097
7	0.319	0.290	0.090	42	0.145	0.125	0.137
8	0.165	0.174	0.057	43	0.268	0.259	0.036
9	0.173	0.201	0.162	44	0.272	0.247	0.091
10	0.298	0.269	0.097	45	0.060	0.046	0.231
11	0.389	0.416	0.069	46	0.298	0.323	0.087
12	0.235	0.284	0.208	47	0.192	0.168	0.126
13	0.225	0.222	0.012	48	0.209	0.209	0.001
14	0.102	0.131	0.282	49	0.361	0.343	0.049
15	0.105	0.141	0.340	50	0.112	0.150	0.344
16	0.152	0.163	0.077	51	0.233	0.193	0.171
17	0.403	0.340	0.157	52	0.226	0.215	0.048
18	0.144	0.154	0.071	53	0.253	0.281	0.110
19	0.108	0.124	0.150	54	0.299	0.271	0.093
20	0.273	0.267	0.025	55	0.056	0.071	0.262
21	0.388	0.404	0.042	56	0.138	0.168	0.218
22	0.371	0.387	0.045	57	0.141	0.167	0.188
23	0.456	0.480	0.052	58	0.109	0.106	0.035
24	0.264	0.247	0.064	59	0.201	0.207	0.028
25	0.348	0.330	0.053	60	0.385	0.425	0.105
26	0.257	0.240	0.065	61	0.314	0.289	0.079
27	0.170	0.136	0.198	62	0.120	0.130	0.089
28	0.184	0.149	0.191	63	0.227	0.201	0.113
29	0.353	0.337	0.044	64	0.125	0.120	0.044
30	0.351	0.343	0.023	65	0.416	0.451	0.084
31	0.296	0.255	0.138	66	0.160	0.195	0.219
32	0.356	0.342	0.039	67	0.281	0.278	0.011
33	0.214	0.233	0.088	68	0.162	0.186	0.149
34	0.215	0.277	0.286	69	0.322	0.277	0.141
35	0.172	0.150	0.132	70	0.195	0.231	0.184
				Average			0.113

10.3389/fpls.2023.1283230

TABLE 11 Impurity ratios of 70 samples.

Sample number	Impurity ratio/%			Sample number	Impurity ratio/%		
	Measured	Estimated	Relative errors		Measured	Estimated	Relative errors
1	0.473	0.499	0.055	36	0.328	0.319	0.026
2	0.237	0.254	0.071	37	0.313	0.322	0.032
3	0.423	0.424	0.004	38	0.091	0.104	0.142
4	0.292	0.298	0.021	39	0.217	0.240	0.103
5	0.303	0.263	0.133	40	0.380	0.381	0.001
6	0.602	0.570	0.053	41	0.241	0.240	0.005
7	0.445	0.445	0.002	42	0.369	0.369	0.000
8	0.372	0.341	0.082	43	0.292	0.282	0.035
9	0.393	0.352	0.104	44	0.328	0.337	0.026
10	0.294	0.280	0.048	45	0.146	0.167	0.148
11	0.529	0.554	0.046	46	0.274	0.316	0.156
12	0.277	0.272	0.018	47	0.310	0.273	0.118
13	0.378	0.388	0.028	48	0.410	0.382	0.068
14	0.206	0.199	0.034	49	0.254	0.269	0.063
15	0.332	0.314	0.055	50	0.320	0.319	0.004
16	0.240	0.217	0.098	51	0.343	0.386	0.124
17	0.482	0.452	0.062	52	0.328	0.325	0.009
18	0.277	0.298	0.073	53	0.385	0.355	0.077
19	0.331	0.317	0.043	54	0.211	0.232	0.102
20	0.274	0.265	0.034	55	0.228	0.248	0.088
21	0.358	0.322	0.102	56	0.420	0.389	0.073
22	0.491	0.470	0.042	57	0.268	0.267	0.007
23	0.417	0.439	0.054	58	0.209	0.200	0.043
24	0.286	0.318	0.110	59	0.239	0.245	0.023
25	0.273	0.241	0.119	60	0.427	0.421	0.014
26	0.316	0.337	0.066	61	0.332	0.320	0.036
27	0.267	0.265	0.006	62	0.319	0.315	0.011
28	0.251	0.272	0.082	63	0.253	0.239	0.054
29	0.208	0.249	0.196	64	0.313	0.339	0.082
30	0.375	0.334	0.109	65	0.500	0.483	0.034
31	0.296	0.347	0.173	66	0.418	0.378	0.095
32	0.229	0.290	0.265	67	0.297	0.320	0.077
33	0.283	0.279	0.014	68	0.299	0.337	0.128
34	0.357	0.332	0.072	69	0.475	0.465	0.021
35	0.350	0.342	0.024	70	0.301	0.302	0.002
				Average			0.65

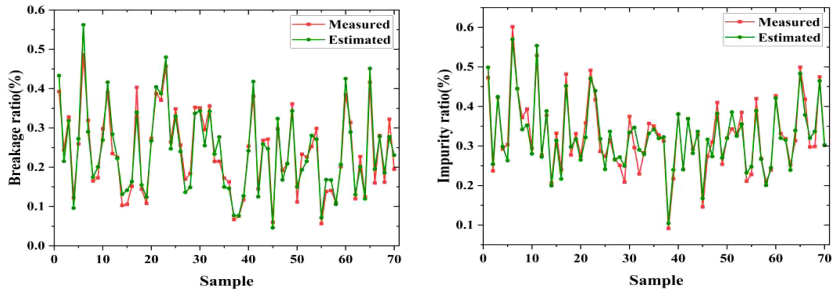


FIGURE 16
Instances of estimation and measured breakage and impurity ratio.

and 0.28. This improvement in accuracy was accomplished with 38.12M, 93.73G, and 10.58ms reduction in Params, FLOPs, and inference time, respectively, making it advantageous for deployment on edge devices and real-time inference.

3. The estimated data obtained according to the approach developed in this study fit the manually obtained breakage and impurity ratios with average relative errors of 11.3% and 6.5%, respectively. The lower segmentation accuracy of broken *cane* is due to their burr and ambiguous boundaries, resulting in a higher average relative error of the breakage ratio.

The raw sugarcane not only includes top and leaf impurities but also contains other impurities like dispersed root whiskers. The upcoming research will emphasize mechanical cleaning of sand, gravel, soil, and similar substances. Additionally, a pivotal aspect of the forthcoming study will involve counting sugarcane roots and estimating their quality through object detection.

Data availability statement

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

Author contributions

XL: Methodology, Writing – original draft, Writing – review & editing. ZZ: Funding acquisition, Methodology, Writing – review & editing. SL: Methodology, Writing – review & editing. TL: Data curation, Investigation, Visualization, Writing – original draft. JZ: Data curation, Visualization, Writing – original draft. TN: Visualization, Writing – original draft. CJ: Investigation, Writing – original draft.

Funding

The author(s) declare financial support was received for the research, authorship, and/or publication of this article. This work is supported in part by earmarked fund for China Agriculture Research System, grant number CARS-17. number CARS-17.

Acknowledgments

We are very grateful to Junshi Sugar Refinery for providing us with the site for collecting experimental data.

Conflict of interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Publisher's note

References

Aparatana, K., Saengprachatanarug, K., Izumikawa, Y., Nakamura, S., and Taira, E. (2020). Development of sugarcane and trash identification system in sugar production using hyperspectral imaging. *J. Infrared Spectrosc.* 28, 133–139. doi: 10.1177/0967033520905369

Chen, J., Lian, Y., and Li, Y. (2020). Real-time grain impurity sensing for rice combine harvesters using image processing and decision-tree algorithm. *Comput. Electron. Agric.* 175, 105591. doi: 10.1016/j.compag.2020.105591

Chen, L.-C., Zhu, Y., Papandreou, G., Schroff, F., and Adam, H. (2018). "Encoder-decoder with atrous separable convolution for semantic image segmentation," in *Proceedings of the European conference on computer vision (ECCV)*. (Germany: ECCV) 801–818. doi: 10.1007/978-3-030-01234-249

Chen, M., Jin, C., Ni, Y., Xu, J., and Yang, T. (2022). Online detection system for wheat machine harvesting impurity rate based on DeepLabV3+. *Sensors* 22 (19), 7627. doi: 10.3390/s22197627

Chinese government website (2018). Available at: http://www.gov.cn/zhengce/content/2018-12/29/content_5353308.htm?ivk_sa=1024320u (Accessed May 24, 2023).

de Mello, M. L., Barros, N. Z., Sperança, M. A., and Pereira, F. M. V. (2022). Impurities in raw sugarcane before and after biorefinery processing. *Food. Anal. Methods* 15, 96–103. doi: 10.1007/s12161-021-02105-1

Dos Santos, L. J., Filletti, É.R., and Pereira, F. M. V. (2021). Artificial intelligence method developed for classifying raw sugarcane in the presence of the solid impurity. *Eclética Quím.* 46 (3), 49–54. doi: 10.26850/1678-4618eqj.v46.3.2021.p49-54

Guedes, W. N., dos Santos, L. J., Filletti, É.R., and Pereira, F. M. V. (2020). Sugarcane stalk content prediction in the presence of a solid impurity using an artificial intelligence method focused on sugar manufacturing. *Food. Anal. Methods* 13, 140–144. doi: 10.1007/s12161-019-01551-2

Guedes, W. N., and Pereira, F. M. V. (2018). Classifying impurity ranges in raw sugarcane using laser-induced breakdown spectroscopy (LIBS) and sum fusion across a tuning parameter window. *Microchem. J.* 143, 331–336. doi: 10.1016/j.microc.2018.08.030

Guedes, W. N., and Pereira, F. M. V. (2019). Raw sugarcane classification in the presence of small solid impurity amounts using a simple and effective digital imaging system. *Comput. Electron. Agric.* 156, 307–311. doi: 10.1016/j.compag.2018.11.039

Hou, Q., Zhang, L., Cheng, M.-M., and Feng, J. (2020). "Strip pooling: Rethinking spatial pooling for scene parsing," in *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*. (Seattle, WA, USA), 4003–4012. doi: 10.1109/CVPR42600.2020.00406

Hou, Q., Zhou, D., and Feng, J. (2021). "Coordinate attention for efficient mobile network design," in *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*. (Nashville, TN, USA), 13713–13722. doi: 10.1109/CVPR46437.2021.01350

Jin, C., Liu, S., Chen, M., Yang, T., and Xu, J. (2022). Online quality detection of machine-harvested soybean based on improved U-Net network. *Trans. Chin. Soc Agric. Eng.* 38, 70–80. doi: 10.11975/j.issn.1002-6819.2022.16.008

Li, T., Tong, J., Liu, M., Yao, M., Xiao, Z., and Li, C. (2022). Online detection of impurities in corn deep-bed drying process utilizing machine vision. *Foods* 11, 4009. doi: 10.3390/foods11244009

Lin, T.-Y., Goyal, P., Girshick, R., He, K., and Dollár, P. (2017). "Focal loss for dense object detection," in *IEEE Transactions on Pattern Analysis and Machine Intelligence*. 42 (2), 318–327. doi: 10.1109/TPAMI.2018.2858826

Lin, P., Li, D., Jia, Y., Chen, Y., Huang, G., Elkhouchlaa, H., et al. (2022). A novel approach for estimating the flowering rate of litchi based on deep learning and UAV images. *Front. Plant Sci.*, 3001. doi: 10.3389/fpls.2022.966639

Liu, L., Du, Y., Chen, D., Li, Y., Li, X., Zhao, X., et al. (2022). Impurity monitoring study for corn kernel harvesting based on machine vision and CPU-Net. *Comput. Electron. Agric.* 202, 107436. doi: 10.1016/j.compag.2022.107436

Liu, Q., Liu, W., Liu, Y., Zhe, T., Ding, B., and Liang, Z. (2023). Rice grains and grain impurity segmentation method based on a deep learning algorithm-NAM-EfficientNetv2. *Comput. Electron. Agric.* 209, 107824. doi: 10.1016/j.compag.2023.107824

Luo, Z., Yang, W., Yuan, Y., Gou, R., and Li, X. (2023). Semantic segmentation of agricultural images: A survey. *Inf. Process. Agric.* doi: 10.1016/j.inpa.2023.02.001

Martins, M. B., and Ruiz, D. (2020). Influence of operational conditions of mechanized harvesting on sugarcane losses and impurities. *Eng. Agric.* 40, 352–355. doi: 10.1590/1809-4430-Eng.Agric.v40n3p352-355/2020

Meng, L., Xu, L., and Guo, J. (2020). Semantic segmentation algorithm based on improved MobileNetv2. *Acta Electonica Sin.* 48, 1769. doi: 10.3969/j.issn.0372-2112.2020.09.015

Milletari, F., Navab, N., and Ahmadi, S.-A. (2016). "V-net: Fully convolutional neural networks for volumetric medical image segmentation," in *2016 Fourth International Conference on 3D Vision (3DV). IEEE.* (Stanford, CA, USA: IEEE) 2016, 565–571. doi: 10.1109/3DV.2016.79

Momin, M. A., Yamamoto, K., Miyamoto, M., Kondo, N., and Grift, T. (2017). Machine vision based soybean quality evaluation. *Comput. Electron. Agric.* 140, 452–460. doi: 10.1016/j.compag.2017.06.023

National Development and Reform Commission (2023). Available at: https://www.ndrc.gov.cn/fgsj/tjsj/jjmy/zyspqk/202302/t20230217_1348905_ext.html (Accessed May 24 2023).

Peng, H., Zhong, J., Liu, H., Li, J., Yao, M., and Zhang, X. (2023). ResDense-focal-DeepLabV3+ enabled litchi branch semantic segmentation for robotic harvesting. *Comput. Electron. Agric.* 206, 107691. doi: 10.1016/j.compag.2023.107691

Rong, D., Wang, H., Xie, L., Ying, Y., and Zhang, Y. (2020). Impurity detection of juglans using deep learning and machine vision. *Comput. Electron. Agric.* 178, 105764. doi: 10.1016/j.compag.2020.105764

Ronneberger, O., Fischer, P., and Brox, T. (2015). "U-net: Convolutional networks for biomedical image segmentation," in *Medical Image Computing and Computer-Assisted Intervention-MICCAI 2015: 18th International Conference, Munich, October 5-9, 2015, Proceedings, Part III 18* (Cham: Springer), 234–241. doi: 10.1007/978-3-319-24574-4_28

Sandler, M., Howard, A., Zhu, M., Zhmoginov, A., and Chen, L.-C. (2018). "Mobilenetv2: Inverted residuals and linear bottlenecks," in *IEEE/CVF Conference on Computer Vision and Pattern Recognition*. (Salt Lake City, UT, USA). 2018, 4510–4520. doi: 10.1109/CVPR.2018.00474

Selvaraju, R. R., Cogswell, M., Das, A., Vedantam, R., Parikh, D., and Batra, D. (2020). Grad-CAM: Visual explanations from deep networks via gradient-based localization. *Int. J. Comput. Vis.* 128, 336–359. doi: 10.1109/ICCV.2017.74

Shen, Y., Yin, Y., Zhao, C., Li, B., Wang, J., Li, G., et al. (2019). Image recognition method based on an improved convolutional neural network to detect impurities in wheat. *IEEE Access* 7, 162206–162218. doi: 10.1109/ACCESS.2019.2946589

Wang, P., Chen, P., Yuan, Y., Liu, D., Huang, Z., Hou, X., et al. (2018). "Understanding convolution for semantic segmentation," in *IEEE Winter Conference on Applications of Computer Vision (WACV)*. (Lake Tahoe, NV, USA), 2018, 1451–1460. doi: 10.1109/WACV.2018.00163

Wu, Z., Yang, R., Gao, F., Wang, W., Fu, L., and Li, R. (2021). Segmentation of abnormal leaves of hydroponic lettuce based on DeepLabV3+ for robotic sorting. *Comput. Electron. Agric.* 190, 106443. doi: 10.1016/j.compag.2021.106443

Wu, F., Yang, Z., Mo, X., Wu, Z., Tang, W., Duan, J., et al. (2023). Detection and counting of banana bunches by integrating deep learning and classic image-processing algorithms. *Comput. Electron. Agric.* 209, 107827. doi: 10.1016/j.compag.2023.107827

Xie, L., Wang, J., Cheng, S., Zeng, B., and Yang, Z. (2018). Optimisation and finite element simulation of the chopping process for chopper sugarcane harvesting. *Biosyst. Eng.* 175, 16–26. doi: 10.1016/j.biosystemseng.2018.08.004

Xie, E., Wang, W., Yu, Z., Anandkumar, A., Alvarez, J. M., and Luo, P. (2021). SegFormer: Simple and efficient design for semantic segmentation with transformers. *rXiv [Preprint]*. doi: 10.48550/arXiv.2105.15203

Xu, T., Ma, A., Lv, H., Dai, Y., Lin, S., and Tan, H. (2023). A lightweight network of near cotton-coloured impurity detection method in raw cotton based on weighted feature fusion. *IET Image Process* 17 (9), ipr2.12788. doi: 10.1049/ipr2.12788

Yu, L., Qian, M., Chen, Q., Sun, F., and Pan, J. (2023). An improved YOLOv5 model: application to mixed impurities detection for walnut kernels. *Foods* 12, 624. doi: 10.3390/foods12030624

Yu, J., Zhang, J., Shu, A., Chen, Y., Chen, J., Yang, Y., et al. (2023). Study of convolutional neural network-based semantic segmentation methods on edge intelligence devices for field agricultural robot navigation line extraction. *Comput. Electron. Agric.* 209, 107811. doi: 10.1016/j.compag.2023.107811

Zhang, C., Li, T., and Li, J. (2022). Detection of impurity rate of machine-picked cotton based on improved canny operator. *Electronics* 11, 974. doi: 10.3390/electronics11070974

Zhao, H., Shi, J., Qi, X., Wang, X., and Jia, J. (2017). "Pyramid scene parsing network," in *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*. (Honolulu, HI, USA). 2017, 2881–2890. doi: 10.1109/CVPR.2017.660

Zhu, S., Ma, W., Lu, J., Ren, B., Wang, C., and Wang, J. (2023). A novel approach for apple leaf disease image segmentation in complex scenes based on two-stage DeepLabv3+ with adaptive loss. *Comput. Electron. Agric.* 204, 107539. doi: 10.1016/j.compag.2022.107539

frontiers | Frontiers in Plant Science

TYPE Original Research
PUBLISHED 08 November 2023
DOI 10.3389/fpls.2023.1247156

OPEN ACCESS

EDITED BY
Xi Tian,
Beijing Academy of Agriculture and Forestry Sciences, China

REVIEWED BY
Geza Bujdoso,
Eszterhazy Karoly Catholic University, Hungary
Tanima Bhattacharya,
Lincoln University College, Malaysia

*CORRESPONDENCE
Fujie Zhang
20030031@kust.edu.cn
Yumin Zeng
yuminzengkm@126.com

RECEIVED 25 June 2023
ACCEPTED 11 October 2023
PUBLISHED 08 November 2023

CITATION
Zhan Z, Li L, Lin Y, Lv Z, Zhang H, Li X, Zhang F and Zeng Y (2023) Rapid and accurate detection of multi-target walnut appearance quality based on the lightweight improved YOLOv5s_AMM model.
Front. Plant Sci. 14:1247156.
doi: 10.3389/fpls.2023.1247156

Rapid and accurate detection of multi-target walnut appearance quality based on the lightweight improved YOLOv5s_AMM model

Zicheng Zhan[1], Lixia Li[1], Yuhao Lin[1], Zhiyuan Lv[1], Hao Zhang[1], Xiaoqing Li[2], Fujie Zhang[1]* and Yumin Zeng[3]*

[1]Laboratory of Physical Properties of Agricultural Materials, College of Modern Agricultural Engineering, Kunming University of Science and Technology, Kunming, Yunnan, China, [2]69223 Troops, People's Liberation Army, Aksu, Xinjiang Uygur Autonomous Region, China, [3]Project Management Division, Yunnan Provincial Forestry and Grassland Technology Extension Station, Kunming, Yunnan, China

Introduction: Nut quality detection is of paramount importance in primary nut processing. When striving to maintain the imperatives of rapid, efficient, and accurate detection, the precision of identifying small-sized nuts can be substantially compromised.

Methods: We introduced an optimized iteration of the YOLOv5s model designed to swiftly and precisely identify both good and bad walnut nuts across multiple targets. The M3-Net network, which is a replacement for the original C3 network in MobileNetV3's YOLOv5s, reduces the weight of the model. We explored the impact of incorporating the attention mechanism at various positions to enhance model performance. Furthermore, we introduced an attentional convolutional adaptive fusion module (Acmix) within the spatial pyramid pooling layer to improve feature extraction. In addition, we replaced the SiLU activation function in the original Conv module with MetaAconC from the CBM module to enhance feature detection in walnut images across different scales.

Results: In comparative trials, the YOLOv5s_AMM model surpassed the standard detection networks, exhibiting an average detection accuracy (mAP) of 80.78%, an increase of 1.81%, while reducing the model size to 20.9 MB (a compression of 22.88%) and achieving a detection speed of 40.42 frames per second. In multi-target walnut detection across various scales, the enhanced model consistently outperformed its predecessor in terms of accuracy, model size, and detection speed. It notably improves the ability to detect multi-target walnut situations, both large and small, while maintaining the accuracy and efficiency.

Discussion: The results underscored the superiority of the YOLOv5s_AMM model, which achieved the highest average detection accuracy (mAP) of 80.78%, while boasting the smallest model size at 20.9 MB and the highest frame rate of 40.42 FPS. Our optimized network excels in the rapid, efficient, and accurate detection of mixed multi-target dry walnut quality, accommodating lightweight edge devices. This research provides valuable insights for the detection of multi-target good and bad walnuts during the walnut processing stage.

KEYWORDS

MobileNetV3, ACMIX, MetaAconC, multi-target, target detection, walnut

1 Introduction

Walnuts (*Juglans* spp.) rank among the world's top four dried fruits, alongside almonds, cashews, and hazelnuts. Two predominant species of walnuts, common walnuts (*Juglans regia*) and dark-grained walnuts (*Juglans sigillata*), are extensively cultivated globally. *Juglans sigillata*, also known as iron walnut or Yunnan walnut, is an endemic species in Southwest China. It is distinguished by superior seed quality, full kernels, high protein and fat content, and rich nutritional value (Xie et al., 2021). After degreening, rinsing, and drying, the evaluation of the appearance quality of walnuts plays a vital role in bolstering their market competitiveness. Yunnan walnuts, which are characterized by uneven kernel surfaces, non-uniform maturity, varying harvest patterns, and irregular fruit sizes, pose challenges during processing. Existing green walnut peeling machines often yield unsatisfactory results, leaving behind impurities, surface contamination, and an increased susceptibility to breakage (Su et al., 2021). In accordance with the "Walnut Nut Quality Grade" standard GBT20398-2021,[1] common external defects in walnuts encompass fractured walnut shells, black spots, and insect holes. Black spots on walnut endocarps typically stem from improper peeling, which leaves a residual walnut pericarp on the surface, leading to oxidation and the formation of black spots. In addition to detracting from the appearance quality and grade, these black patches cause mildew due to their moisture-absorbing properties. Furthermore, damaged and insect-infested walnuts expose their kernels to external elements, resulting in rapid deterioration, mould formation, and potentially hazardous substances, such as aflatoxins, due to water infiltration during cleaning. Consequently, there is an urgent need for a rapid and precise method to identify these external defects during walnut production and processing (Li et al., 2019).

Currently, two main approaches are employed to assess produce quality: destructive and non-destructive methods. Destructive methods are utilized to determine the physicochemical or biochemical properties of the produce but require the complete annihilation of the tested specimens, imposing strict technical prerequisites. Although they provide additional phenotypic data, their inherent delay in detection is a drawback. By contrast, non-destructive methods offer advantages such as reduced costs, heightened detection accuracy, and the ability to evaluate produce without damaging it (Arunkumar et al., 2021). Both domestic and international scholars have extensively investigated various non-destructive testing methods for fruits and nuts, including X-ray techniques, acoustic methods (Cobus and van Wijk, 2023), and machine vision approaches (Chakraborty et al., 2023). However, it is worth noting that although these methods excel in detection accuracy, X-ray detection can be expensive, and acoustic methods may be limited to single-target fruit detection, potentially restricting their applicability to primary processing firms.

Deep learning, a non-destructive approach, can swiftly detect issues in one or two phases, offering precise detection and quality control for all types of nuts through computer vision technologies and robotics. The integration of deep learning technology can significantly enhance the production efficiency and quality management within nut processing enterprises by refining the classification and grading processes, automating quality management procedures, and effectively identifying nut defects and abnormalities.

In recent years, researchers have explored a two-stage deep learning approach for fruit and nut detection. For instance, Rika Sustika et al. (2018) investigated the impact of various deep convolutional neural network structures (AlexNet, MobileNet, GoogLeNet, and Xception) on the accuracy of a strawberry grading system (appearance quality detection), with VGGNet demonstrating the highest accuracy (Sustika et al., 2018). Costa et al. (2021) combined machine vision techniques with the Mask-RCNN algorithm (Costa et al., 2021) to detect and semantically segment pecan peel and hull. Fan et al. (2021) proposed an improved rapid R-CNN algorithm (Fan et al., 2021) for the precise detection of green pecans in natural environments. The enhancements included batch normalization, an improved RPN with bilinear interpolation, and the integration of a hybrid loss function. For robot recognition and the picking of walnuts in complex environments, the model achieved an accuracy of 97.71%, a recall rate of 94.58%, an F1 value of 96.12%, and faster detection times. However, these two-stage approaches, which are capable of high accuracy, tend to have slower detection speeds and require lengthy training periods, making them challenging to implement in actual industrial production settings. By contrast, the one-stage approach, represented by the YOLO series algorithm, offers advantages such as fast real-time detection, high accuracy, and robustness. Hao et al. (2022) used an improved YOLOv3 deep learning method for the real-time detection of green walnuts in a natural environment. They pre-trained the model network using the COCO dataset, optimized the performance with data augmentation and K-means clustering, and selected the MobileNetV3 backbone for high accuracy and rapid detection. This approach achieved an average accuracy (mAP) of 86.11% and provided technical support for intelligent orchard management and yield estimation of walnut orchards (Hao et al., 2022). Recognizing the widespread acceptance of the YOLOv5 model as a faster, more accurate, and more efficient target detection model, Yu et al. (2023) proposed an improved walnut kernel impurity detection model based on the YOLOv5 network model. Their model included a small target detection layer, a transformer-encoder module, a convolutional block attention module, and a GhostNet module, leading to enhanced recognition accuracy for small and medium impurities in pecan kernels (Yu et al., 2023). In general, the two-stage target detection algorithm applied to pecans struggles to balance recognition accuracy and detection speed. On the other hand, one-stage algorithm research is tailored to green walnut recognition scenarios. These models and methods cannot be directly applied to the recognition of good- and bad-dried walnuts because of differences in image datasets, such as variations in field orchard backgrounds and occlusion, lighting conditions, or diffractive indices. Additionally, there is limited research on target detection based on deep learning for the classification of dried walnut quality after degreening, washing, and drying during the initial processing stage. Detecting dried walnuts of various target sizes within a wide field of vision

1 https://openstd.samr.gov.cn/bzgk/gb/newGbInfo?hcno=11E65C73CF8B9E071CE76716628F2F80]

presents a challenging task. Therefore, achieving efficient and precise sorting of good and bad dried walnuts using deep learning models has become an urgent matter, significantly impacting the advancement of the entire walnut primary processing industry.

This study introduced an enhanced YOLOv5s_AMM multi-target sorting model tailored for walnuts. (1) The M3-Net network replaced YOLOv5s' C3 structure with MobileNetV3. This substitution has advantages, such as faster inference, heightened accuracy, reduced memory usage, and improved feature representation. Consequently, it emerges as a superior option for target detection in devices with resource constraints. (2) The model achieved enhanced classification accuracy, adaptive mixture modelling, rapid training and inference, and robustness against noise and outliers by incorporating the novel ACMIX paradigm. ACMIX integrates convolution with self-attentiveness after an SPP (spatial pyramid pooling) layer (He et al., 2015). (3) In the neck layer, the CBM module replaced the activation function of the conventional Conv2d convolution layer with MetaAconC. This substitution results in performance enhancements, adaptive activation, non-linear and smoothing behavior, computational efficiency, and robustness against noise and outliers. Finally, the improved YOLOv5s_AMM detection model, when practically applied to differentiate between good and bad walnuts of various sizes, achieved real-time and efficient classification. This advancement has significant practical value for enhancing walnut detection efficiency, quality, and market competitiveness. This is particularly beneficial to primary processing enterprises aiming to increase the value of their walnut products and contributes to the growth of a more intelligent and integrated walnut industry.

2 Materials and methods

2.1 Image sample acquisition

In this study, walnuts were sourced from Fengqing County, Lincang City, Yunnan Province, China. The RGB images used for the analysis were collected at the Agricultural Material Characterization Laboratory at the Kunming University of Technology. These images were captured from 9:00 a.m. to 6:00 p.m. on December 24–26, 2022. For image acquisition, we employed a Hikvision industrial camera (model MV-CA050-20GC) with a 5-megapixel resolution and a CMOS Gigabit Ethernet industrial surface array camera capable of producing images at a resolution of 2592×2048. The images were saved in the JPEG format. The camera was securely mounted at a height of 195 **cm** above the ground and positioned 95 **cm** above the surface level using an adjustable aluminium mount. All images were captured under consistent conditions, including the same camera height, uniform light source brightness, consistent background, and roller guide profile phase. The image capture date was December 24, 2022. During the image capture process, we used an exposure time of 4,000 μs and frame rate of 1 in the continuous mode of the camera. This setup allowed us to capture images of walnuts in their natural state, as depicted in Figure 1, with the walnuts evenly distributed on the moving part of the profiled roller-wheel guide.

According to the national standard classification GB/T20398-2021 for walnut quality grades, our evaluation considered various factors such as walnut uniformity, shell integrity, color, and suture line tightness. Based on these criteria, we classified walnuts into two categories: (1) good walnuts (Figure 2A), characterized by intact shells and primarily exhibiting a yellow-white color, and (2) bad walnuts (Figures 2B, C), including walnuts with black spots (Figure 2B) and walnuts with broken shells (Figure 2C). In this study, 120 original images with a resolution of 2,592×2,048 were acquired, and multi-target walnuts with excellent walnuts and bad walnuts (black spots and broken) were randomly inserted into this dataset.

2.2 Dataset construction

During the data processing phase, we divided the initial dataset of 120 images, each with a resolution of 2,592×2,048, into 2,000 images with a resolution of 640×640. We employed LabelImg software for manual annotation, marking the location boxes, and categorizing walnuts as either good or bad within the original images. This annotation process produced corresponding annotation files. Upon completing the image annotation, we randomly divided the entire dataset into three sets: training, validation, and test. The distribution ratio was 8:1:1, ensuring adequate data for training and model evaluation. In statistical terms, each image in this study contained 5–40 walnuts, resulting in a total of 53,301 labels within the walnut dataset. Among these labels, 25,099 were associated with good walnuts, whereas 28,208 were assigned to bad walnuts. This distribution indicates a reasonably balanced dataset, with a ratio of approximately 0.88 between the two image categories.

Before commencing model training, we subjected the walnut training set to a combination of offline data enhancement techniques, including contrast adjustment, scaling, luminance modification, pretzel noise, and Gaussian noise (Taylor and Nitschke, 2018). These techniques were applied randomly. As shown in Figure 3, They encompassed four specific methods; (1)

FIGURE 1
Map of the walnut image acquisition environment.

FIGURE 2
Types of walnuts. **(A–C)** Good walnuts **(A)** and defective walnuts (black spots [**B**] and broken [**C**]).

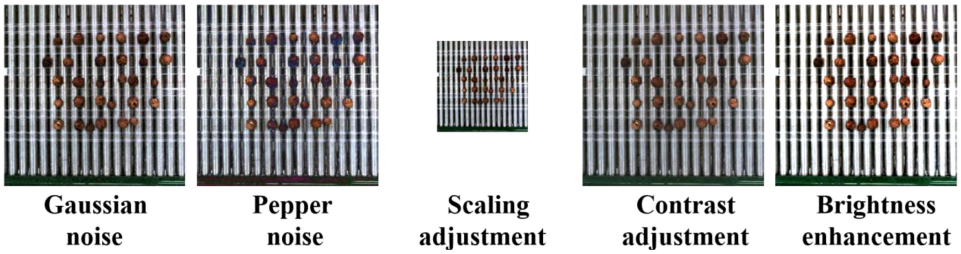


FIGURE 3
Schematic of the image enhancement of the walnut dataset.

random contrast enhancement within the range of 0.7 to 1.4 with a gradient of 0.05; (2) random scaling of the training set within the range of 0.5 to 1.5 with a gradient of 0.1; (3) random luminance adjustment for the training set within the range of 0.6 to 1.4; and (4) random modification of the training set's luminance to either 50–150% of random Gaussian noise or random pretzel noise within the same range. As shown in Table 1, these data enhancement procedures resulted in a 25% increase in the number of training sets. Consequently, the walnut dataset contained a total of 2,000 image data entries after data enhancement.

2.3 YOLOv5 network model and optimization structure

2.3.1 YOLOv5 model

The YOLO Network Series provides a rapid and efficient solution for real-time target detection tasks, delivering high accuracy and the capability to identify targets of various sizes. Its versatility extends to a wide range of applications, including autonomous vehicles, robotics, and surveillance systems.

TABLE 1 Classification of the walnut image dataset.

Dataset	Original image	Enhanced image
Training set	1200	1598
Validation set	150	201
Test set	150	201
Total	1500	2000

YOLOv5 outperforms YOLOv3 and YOLOv4 in terms of rapid and precise real-time target detection. It achieves this superiority while maintaining a lighter, more efficient, and more easily deployable profile on resource-constrained devices. These improvements encompass several key aspects. (1) Enhanced backbone network: YOLOv5 adopts the CSP-Darknet53 architecture (Bochkovskiy et al., 2020) as its backbone network. This innovation improves feature extraction while reducing the computational cost. (2) Innovative neck layers: The model incorporates SPP and PAN neck layers. These layers combine features from different layers, thereby enhancing both the accuracy and efficiency of the model. (3) Optimized training process: YOLOv5 benefits from the optimized training process. This includes a new hyperparameter search algorithm that efficiently tunes model settings, a novel loss function, improved data augmentation techniques, and AutoAugment, which automatically identifies the optimal hyperparameters. The amalgamation of these enhancements enables YOLOv5 to deliver cutting-edge target detection performance while maintaining a real-time processing speed.

The official code allows for the training of four distinct object detection models with varying depths and widths. In the YOLOv5 series, YOLOv5s serves as the baseline with the smallest depth and width. The other three networks build on this foundation, becoming deeper and more complex. These networks incorporate additional

convolutional layers and residual modules in the backbone and employ more channels in the head module to enhance accuracy.

Table 2 provides a comparison of the accuracy, model size, and detection performance across the four distinct YOLOv5 models. In terms of detection accuracy, YOLOv5s exhibited a slightly lower mAP than YOLOv5m (1.67% lower), YOLOv5l (3.47% lower), and YOLOv5x (3.36% lower). However, when considering the model size, YOLOv5s stood out because of its compact size of 27.1 MB, which was notably smaller than YOLOv5m, offering a reduction of 53.5 MB. This size advantage makes YOLOv5s a cost-effective choice, particularly for deployment on embedded devices, where storage constraints are critical. In terms of detection speed, YOLOv5s outperformed the other models, detecting 7.85 frames per second more than YOLOv5m, 17.13 frames more than YOLOv5l, and 25.97 frames more than YOLOv5x. This superior inference speed position of YOLOv5s is an excellent option for real-time detection scenarios and applications demanding rapid responses. Given the emphasis on low latency and cost-effective deployment for lightweight multi-target kernel peach detection, YOLOv5s presented a compelling proposition with a detection accuracy of 78.97%, a model size of 27 MB, and a detection speed of 47 FPS. It effectively balances accuracy, model size, and inference speed, making it a well-suited base model for further enhancement.

2.3.2 MobileNetV3: lightweight backbone network

The concept of a lightweight backbone pertains to neural network architectures optimized for target detection tasks. This optimization involves a reduction in the number of network parameters and layers, while maintaining high accuracy in the target detection tasks. The core objective was to curtail the computational burden and memory requirements of the network. The integration of lightweight backbone networks into target detection models yields substantial advantages, including enhanced computational efficiency, reduced memory demand, accelerated inference speed, and increased robustness. Consequently, they have gained popularity, particularly for resource-constrained applications. In this context, Andrew Howard et al. (2019) proposed the "MobileNetV3" architecture in their research titled "Searching for MobileNetV3" (Howard et al., 2019). Their work demonstrated that MobileNetV3 outperforms alternatives such as ShuffleNet (Bhattacharya et al., 2006) and MobileNetV2 (Sandler et al., 2018) in terms of accuracy, advanced features, and efficient training. This renders MobileNetV3 a versatile and effective option, particularly for resource-constrained devices. Figure 4 illustrates the structure of MobileNetV3, which includes 1×1 convolutional layers to adjust channel numbers, deep convolutions in high-dimensional spaces, SE attention mechanisms for feature map optimization, and 1×1 convolutional layers for channel number reduction (employing linear activation functions). The network employs residual connections when the step size is 1, and the input and output feature maps have the same shape, whereas in the downsampling stage (step size = 2), the downsampled feature maps are directly output. MobileNetV3's architectural contributions are primarily grouped into the following categories:

TABLE 2 Comparison of the prediction results from YOLOv5 models.

Model	mAP@0.5 (%)	Parameters	Model size (MB)	FPS
YOLOv5s	78.97	7,276,605	27.1	47.22
YOLOv5m	80.64	7,276,605	80.6	39.37
YOLOv5l	82.44	7,276,605	178	30.09
YOLOv5x	82.33	7,276,605	333	21.25

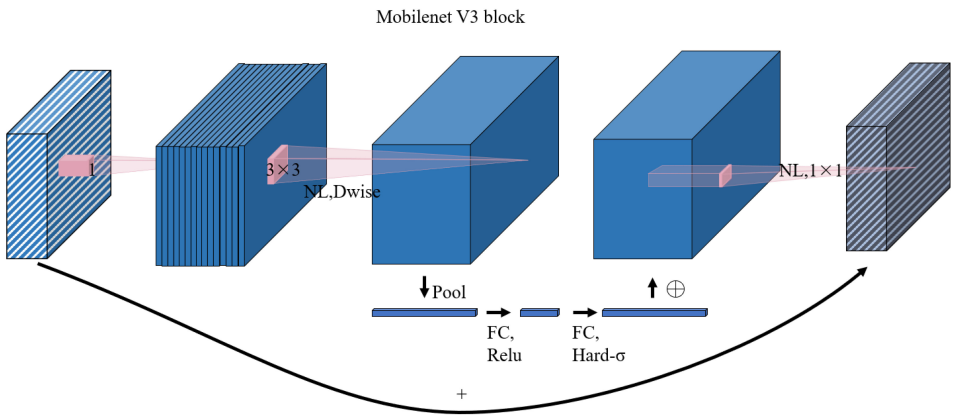


FIGURE 4
Structural diagram of the MobileNetV3 network.

(1) MobileNetV3 leverages deeply separable convolutions and residual blocks to reduce parameters and computations, thereby enhancing the computational efficiency.

(2) Fewer layers are used to minimize the memory requirements and facilitate deployment on resource-constrained devices.

(3) MobileNetV3 incorporates an SE attention mechanism and a hard-swish activation function to support data representation capabilities, while maintaining a modest parameter count and computational load.

(4) The utilization of hybrid precision training and knowledge distillation techniques further enhances the training effectiveness while reducing memory and computational costs.

MobileNetV3 attains state-of-the-art performance across various tasks while retaining its lightweight and efficient nature. This results in substantial reductions in computational and memory costs, rendering it an ideal choice for target detection in resource-constrained devices. This study refined the MobileNetV3 model to enhance its suitability as a lightweight backbone network, thereby achieving higher accuracy and improved network performance.

2.3.3 Acmix: attention-based convolutional hybrid structure

The Acmix architecture (attention-convolution hybrid), introduced in 2021, represents a novel neural network architecture comprising primarily three fundamental modules: an attention module, a convolution module, and a hybrid module. The attention module is responsible for capturing essential features within the input image. Both global and local attention modules are utilized in the Acmix architecture. The global attention module captures the image's broader contextual information, whereas the local attention module focuses on capturing intricate details within the image. The primary function of the convolution module is the feature extraction from the input image. To achieve this, the Acmix architecture combines the conventional convolutional layers with depth-separable convolutional layers. This integration significantly reduces the computational cost and memory requirements of the network, thereby enhancing its overall efficiency. The hybrid module serves as the nexus where the features extracted by the attention module and convolution module converge and interact. In this context, the Acmix architecture uses both global and local hybrid modules. The global hybrid module harmonizes characteristics from the global attention module and convolutional module, whereas the local hybrid module fuses attributes from the local attention module and convolutional module.

Figure 5 illustrates the hybrid module proposed by Acmix. The left diagram shows the flowchart of the conventional convolution and self-attention module. (a) The output of the 3×3 convolutional layer can be decomposed into a summation of shifted feature maps. Each of these feature maps was generated by applying a 1×1 convolution with kernel weights at specific positions, denoted by s(x,y). (b) The self-attention process involves projecting the input feature maps into queries, keys, and values, followed by 1×1 convolution. The attention weights computed through the query-key interaction were used to aggregate the values. Conversely, the diagram on the right delineates the pipeline of our module. (c) Acmix operates in two stages. In stage one, the input feature map underwent projection using three 1×1 convolutions. Stage two employs intermediate features using two examples. The characteristics extracted from both paths are fused to generate the final output. The computational complexity of each operation block is shown in the upper corner (Pan et al., 2022). The Acmix architecture has demonstrated state-of-the-art performance across various benchmark datasets for image classification tasks, while maintaining a lightweight and efficient design. The attention and hybrid modules within Acmix are strategically designed to capture both global and local features within walnut images, with a particular emphasis on identifying black spots and damaged areas on walnuts. Consequently, the Acmix module is introduced after the SPP module during the feature fusion phase of the improved model to enhance its performance, particularly on intricate datasets.

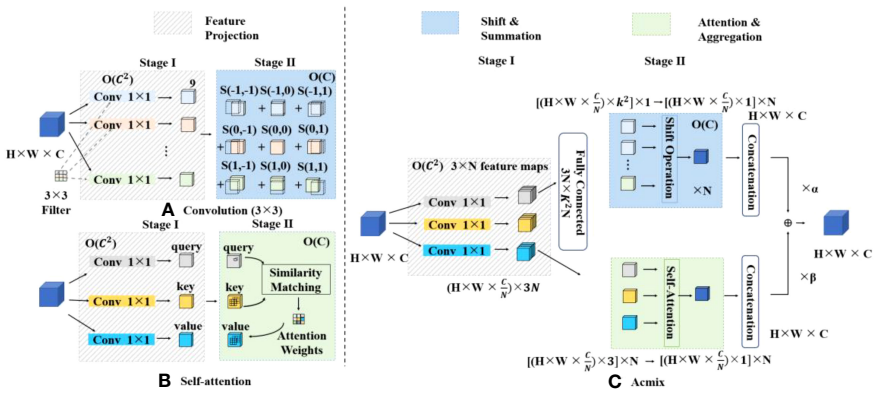


FIGURE 5
Structure of the hybrid module network in Acmix.

2.3.4 MetaAconC activation function

$$\text{MetaAconC} = (p_1 - p_2)X\delta[\beta_c(p_1 - p_2)X] + p_2X, \beta_c$$
$$= \delta W_1 W_2 \sum_{h=1}^{H}\sum_{w=1}^{W} X_{c,h,w} \quad (1)$$

In Eq. (1), X represents the input feature map, where X (c,h,w) denotes the feature input with dimensions of C × H × W. W1 and W2 represent the computed weights; p1 and p2 represent adjustable learning parameters; β signifies the adaptive function; and δ represents the sigmoid activation function (Nan et al., 2023). MetaAconC (Ma et al., 2021) is a novel activation function proposed in 2021 to address the limitations of conventional activation functions. This is achieved by combining the Meta-AC and CAN functions, which are known to be vulnerable to the gradient vanishing problem and can lead to neuron inactivity. The Meta-AC function is capable of concatenating multiple activation functions and adapting to the specific input data distribution. The CAN function non-linearly transforms the output generated by the Meta-AC function, enabling it to capture more intricate and abstract features. Empirical evidence has shown the superiority of the MetaAconC activation function over traditional alternatives, such as ReLU and sigmoid. It offers distinct advantages, including adaptivity, computational efficiency, and robustness against noise and outliers. These attributes were substantiated in subsequent ablation experiments. In the context of our enhanced model, the original SiLU activation function was replaced by the MetaAconC activation function. The experimental data underscore its suitability for walnut image detection.

2.3.5 Improved Yolov5s network structure

In this study, we built upon the architecture of YOLOv5s, version 5.0, as the foundation for model improvement. The objective was to address issues related to accuracy, model size, and detection speed to develop a more appropriate model for the detection of good and bad walnut fruit targets during the primary processing stage. The overall enhanced network structure is shown in Figure 6. In this model refinement, we opted to replace the original focus layer with CBH and the C3 backbone network structure with MobileNetV3 from the M3-Net network. This alteration was made with the aim of reducing the model size and ensuring a lightweight and efficient design. Furthermore, we introduced the attention convolutional hybrid (Acmix) structure into the neck layer. This addition reduced the computational cost and memory requirements of the network. The attention and hybrid modules within the Acmix architecture are strategically designed to capture both global and local image features, thereby enhancing the model's performance on complex datasets. Finally, we replaced the two Conv2d modules in the neck layer with the CBM modules. In addition, the SiLU activation function found in the original Conv layer was substituted with the MetaAconC activation function. This adjustment is implemented to improve the input-specific data distribution for tuning, ultimately enhancing the feature detection across different image scales.

Table 3 provides an overview of the replacement lightweight backbone network used in this study, with a primary focus on the incorporation of the M3-Net network to construct the backbone network of the enhanced model. Table 3 presents detailed information on various parameters and components. Specifically, "Input" represents the features of the input layer feature matrix; "#Out" represents the number of channels in the output layer feature matrix; "S" represents the step size of the DW convolution; "exp size" represents the size of the first up-convolution; "SE" indicates whether the attention mechanism is employed; "NL" represents the activation function used; "HS" represents the hard-swish activation function; and "RE" represents the ReLU activation function. Within the modified backbone layer of M3-Net, there are primarily four types of MobileNet_Block:

MNB1_1: CBH + SE attention mechanism + CB
MNB1_3: CBR + SE attention mechanism + CB
MNB2_1_1: CBH + CB + SE attention mechanism + HCB
MNB2_2_4: CBR + CBR + CB.

First, the original focus layer was replaced with the Conv_Bn_Bswish layer, resulting in improved model accuracy,

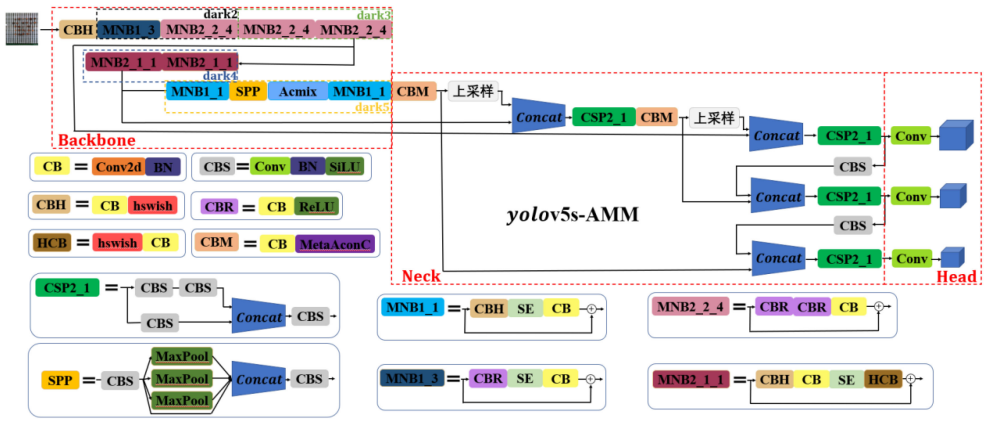


FIGURE 6
Overall network structure of the improved YOLOv5s-AMM model.

TABLE 3 Backbone network with the improved model.

Input	Operator	Exp size	#out	SE	NL	S
640^2×32	Conv_Bn_Hswish	–	64	–	–	1
320^2×64	MobileNet_Block,3×3	64	64	√	RE	2
160^2×64	MobileNet_Block,3×3	384	128	–	RE	2
160^2×128	MobileNet_Block,3×3	448	128	–	RE	1
80^2×128	MobileNet_Block,5×5	512	256	√	HS	2
40^2×256	MobileNet_Block,5×5	960	256	√	HS	2
40^2×256	MobileNet_Block,5×5	512	512	√	HS	1
20^2×512	MobileNet_Block,5×5	512	512	√	HS	2
20^2×512	SPP	–	512	–	–	–
20^2×256	Acmix	–	512	–	–	1
20^2×256	MobileNet_Block,5×5	512	256	√	HS	1

accelerated inference, and architectural simplification. Moreover, the Conv and C3 components of the original dark2 layer were replaced with MNB1_3 and MNB2_2_4, while the Conv and C3 components of the dark3 layer were replaced with two consecutive MNB2_2_4 structures. This replacement strategy employed MobileNetV3_Block to construct a lightweight backbone network, which not only enhanced the accuracy in identifying the walnut dataset but also boosted the network performance efficiency, facilitating faster convergence and superior generalization effects. Subsequently, we introduced the Acmix structure after applying the SPP structure to the output of the final layer. This involved a combination of standard convolutional layers and deeply separable convolutional layers to reduce the neocortex size. Consequently, network performance efficiency was further enhanced, leading to faster convergence. As a result, when the input image size was set to 640×640, the improved backbone network generated output feature maps with dimensions of (20 × 20×1,024), (40×40×512), and (80×80×256). The role of the neck is to integrate the walnut characteristics extracted from the backbone into a format suitable for object detection. This component plays a pivotal role in improving the accuracy of the walnut target detection model by capturing the walnut features at various scales and combining them effectively. This enhances the model's ability to detect objects of varying sizes and aspect ratios. In addition, when replacing the corresponding conv2d module with a CBM module after the (20×20×1024) and (40×40×512) walnut feature maps, the MetaAconC activation function in the CBM module surpasses the performance of the sigmoid function. It is not only adaptive but also capable of learning based on the specific walnut data distribution. This is in contrast with the traditional sigmoid function, which remains fixed and unalterable during the training process. The MetaAconC function offers high computational efficiency, leading to an improved detection performance for walnut images at various scales. Furthermore, it reduces the computational load and memory usage of the network, resulting in shorter inference times and reduced hardware requirements. Finally, the head layer produced output feature maps with dimensions of 80×80×(3×(num_classes+5)), 40×40×(3×(num_classes+5)), and 20×20×(3×(num_classes+5)). Here, "num_classes" denotes the number of detected walnut object classes in the training network, and "3" denotes the number of anchor boxes used for walnut object detection within each grid cell.

2.3.6 Training the multi-target detection model for walnuts (good and bad fruits)

To impart the model with more relevant and informative features, the initial image was segmented into 640×640 pixels, aligned with the model's input image size of 640×640 pixels. Building on this foundation, the model was enhanced using the proposed improvement methodology. Subsequently, the labelled walnut dataset was employed for training within the PyTorch deep learning framework, whereas the validation dataset served as a means to evaluate the effectiveness and performance of the model training process.

Table 4 lists the experimental settings used in this study. Initially, the dataset containing annotations in the VOC format was converted into a format compatible with the YOLOv5 model. Subsequently, the parameters governing the training procedure were configured meticulously. The enhanced YOLOv5s detection network was then subjected to training with an initial learning rate of 0.01, eta_min at 1×10^{-4}, last_epoch at −1, momentum

TABLE 4 Experimental settings for this study.

Name	Value
CPU	AMD Ryzen 9 5900HX with Radeon Graphics octa-core
Memory	32 GB
Storage SSD	1024GB
Graphics card	NVIDIA GeForce RTX 3080
Graphics memory	16GB
Operating System	Windows11
CUDA version	11.6
PyTorch version	1.8.0

10.3389/fpls.2023.1247156

parameter at 0.937, delay parameter at 5×10^{-3}, batch size at 8, and T_max at 250. The optimization during the training procedure was executed using an SGD optimizer. Multi-threaded model training harnessed the computational power of the four processors, whereas the cosine annealing learning rate was dynamically updated for optimization during training. Furthermore, four offline enhancement techniques, including contrast adjustment, scaling, luminance modification, and the introduction of pepper and Gaussian noise to the walnut image data, enrich the contextual information for detecting walnut objects. These augmentations enhance the perception of distinguishing between good and bad walnut features, thereby bolstering the model's robustness and generalization capabilities. Notably, the data augmentation network required approximately 6 h and 52 min to complete the training process.

The entire training process was segmented into two distinct phases, namely, the "freezing phase" and the "thawing phase", in alignment with the underlying model structure. During the freezing phase, the spine of the model remained unaltered and was held constant. No modifications were made to the ad hoc extraction network. During this phase, the focus was on training the weight parameters of the prediction network until they reached a state of saturation and convergence. Subsequently, the model entered the thawing phase, wherein the core of the model was no longer constrained and the weight parameters of the feature extraction network were subjected to training to optimize the entire set of network weights.

The loss curve in target detection serves as a crucial indicator of the training progress of the model by monitoring the value of the LOSS function. This function is a mathematical construct that quantifies the disparity between the model's predicted output, given an input image, and the actual output (i.e., the true value). Within the context of YOLOv5, the loss function comprises several integral components, including localization, confidence, and class loss. The localization loss is responsible for assessing the precision with which the model predicts the coordinates of bounding boxes around objects within an image, whereas the confidence loss quantifies the level of confidence in the model's prediction. The class loss measures the capacity of a model to classify images.

The loss curve indicates the model's learning progress in generating accurate predictions after being trained on the dataset. As the model acquires knowledge from the training data, the loss progressively diminishes. The objective of training is to minimize this loss, which indicates that the model makes accurate predictions on the training data.

As depicted in Figure 7, the initial model observed a notable reduction in the loss value during the fifth iteration. Subsequently, from the fourth to the tenth iteration, the loss value stabilized, hovering at approximately 0.32. Notably, there was no discernible alteration in the thawing stage, even as the model progressed to the 50th iteration.

Following the unfreezing of the model, the loss decreased notably between the 50th and 53rd iterations,.from 0.36 to 0.29. Moreover, between the 60th and 210th iterations, the loss remained consistently lower than that of the original model. Subsequently, a comparison between the original and improved models' loss values from iteration 210 onwards revealed that the improved model exhibited a swifter decline in losses between iterations 210 and

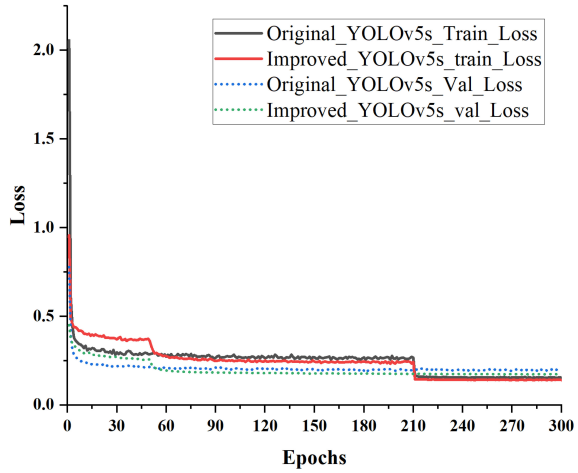


FIGURE 7
Comparison of epochs trained using original and improved models.

300. Ultimately, the loss values of the improved model stabilized at approximately 0.142, a reduction of 0.01 points compared with the original model. These results highlight the superior performance of the enhanced model in distinguishing between good and bad walnuts during the convergence.

3 Results

3.1 Model evaluation indicators

To conduct a thorough evaluation of the model's performance on multi-target walnut images, we employed eight widely accepted evaluation metrics that are commonly used in classical target detection algorithms. These metrics included precision (P), recall (R), F1 score, average precision (AP), average accuracy (mAP), network parameters, model size, and detection speed. Throughout the experimental period, an IoU value of 0.50 was used. To assess real-time detection performance, this study employed frames per second (FPS) as the key metric. A higher FPS indicates a higher model detection rate. Equations (1)–(5) illustrate the specific formulas for calculating P, R, F1, AP, and mAP.

$$P = \frac{TP}{TP + FP} \tag{2}$$

$$R = \frac{TP}{TP + FN} \tag{3}$$

$$F1 - score = \frac{2PR}{P + R} \tag{4}$$

$$AP(k) = \int_0^1 P(R)dR \tag{5}$$

$$mAP = \frac{\sum_1^N AP(k)}{N} \tag{6}$$

where TP represents the number of correctly identified walnuts (true positives); FP represents the instances in which the classifier incorrectly predicted positive samples among the actual negative samples (indicating the number of false negative samples); TN represents the number of correctly identified negative samples; and FN represents the number of negative samples that were incorrectly predicted by the classifier.

The F1 value serves as a comprehensive measure of the overall accuracy of the detection model and is calculated as the average sum of precision and recall. AP represents a measure of the precision and recall trade-off for a given detection model by calculating the area under the recall curve. Higher AP values indicate a better performance. In equation (6), "N" represents the number of object categories, "AP (k)" is the average precision for a specific category (in this study, k=2), and "Σ" signifies the sum across all categories. mAP consolidates accuracy and recall across multiple object categories, offering a global assessment of the object detection model's performance. The scores range from 0 to 1, with higher scores indicating superior performance. Given the need to evaluate an integrated object detection network with multiple object categories and the superiority of the mAP over the F1 score, we chose to use the mAP score for our assessment.

3.2 Experimental results

Table 5 presents the evaluation results of the enhanced model using the 201-objective walnut test dataset. The empirical findings revealed that the improved YOLOv5s model achieved an overall mAP of 80.78% on the test dataset. Additionally, it attained an F1 score of 0.77, a model size of 20.90 MB, and an average detection rate of 40.42 frames per second, thus satisfying the real-time detection requirements. The precision-recall gap across each category ranged from 2.62 to 4.46%. Furthermore, the cumulative mAP for both excellent and bad walnuts was 80.78%. In summary, the enhanced model proposed in this study for the detection of good and bad walnut fruits demonstrates superior accuracy, minimal computational overhead, and rapid inference capabilities.

3.3 Effect of the detection performance of attention mechanisms at various positions

In the context of the original YOLOv5s network, we introduced the attentional convolutional hybrid Acmix module into both the backbone and neck layers, as depicted in Figure 8, to explore the

TABLE 5 Experimental results.

Class	P/%	R/%	mAP@0.5 (%)	F1-score	Model size (MB)	FPS
Good	75.11	70.65	86.79	0.73	20.9	40.42
Bad	82.13	79.51	74.78	0.81		
All	78.62	75.08	80.78	0.77		

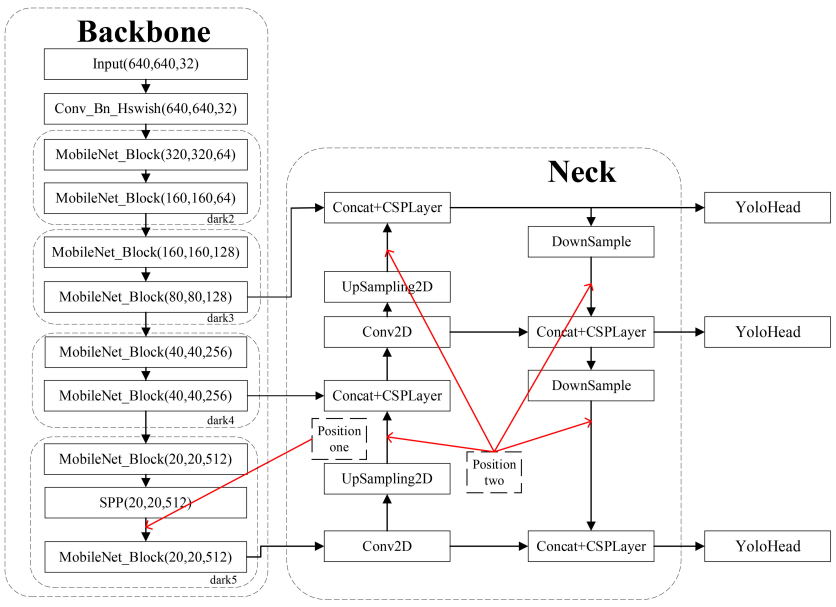


FIGURE 8
Integration of attention mechanisms at different locations in the original YOlOv5s model.

impact of integrating the attentional mechanism at different locations within the model. Specifically, after pyramidal pooling in the SPP space, the final dark5 module at location 1 in the backbone layer incorporated the Acmix module. The Acmix module received the output feature maps from the SPP layer and calculated the channel weights, which were subsequently applied to the input feature maps. This process emphasized significant regions while suppressing insignificant ones. Additionally, at position 2, located after each upsampling and downsampling operation in the neck layer, another Acmix module was applied.

Table 5 presents the results of comparing the effects of adding the attention mechanism at various locations. Notably, the addition of Acmix at location 1 improved the model's mAP by 1.38%, increased the model size by a mere 3.2 MB, and reduced the detection speed by only 4.08 frames per second. However, when added at location 2, the model's mAP experienced a marginal decrease of 0.07%, accompanied by a more significant reduction in detection speed by 18.83 frames per second. This suggests that the Acmix attention mechanism may not be universally applicable to all layers. The reason for this discrepancy lies in the fact that introducing too many attention mechanisms in location 2 of the neck layer may diminish the model's mAP. In addition, the excessive incorporation of attention mechanisms at location 2 introduces a surplus of additional parameters, potentially resulting in network overfitting and deterioration in network performance. The most pronounced enhancement was observed when Acmix was added at location 1, particularly when compared with location 2, or when both locations received attention mechanisms. This is due to the multifaceted scale features generated by SPP, which aid in the detection of targets of varying sizes but have weaker interrelationships. Adding the Acmix attention mechanism after SPP explicitly constructs relationships between features of different scales, enabling the network to better leverage these features and subsequently enhance its performance. Additionally, because SPP generates a surplus of features, some of which may be irrelevant, the Acmix attention mechanism effectively filters out these irrelevant features, focusing on the most pertinent features to reduce feature redundancy. Furthermore, it enables the learning of novel feature expressions based on the features generated by SPP, thereby improving the overall feature representation. Incorporating Acmix attention at position 1 effectively compensated for the limitations of SPP and resulted in more potent feature expressions.

3.4 Effect of various attention mechanisms on detection performance

In the context of the original YOLOv5s model, we introduced various attention mechanisms from position 1 in Figure 8 to investigate their impact on the performance of the target detection model. As shown in Table 6, the addition of the Acmix attention mechanism exhibited the most substantial improvement in the mAP performance compared with the original model, achieving a notable increase of 1.38%. By contrast, the ECA, CBAM, and SE attention mechanisms displayed comparatively less improvement in the mAP performance. This observation underscores that merely applying an attention mechanism after SPP does not inherently improve model accuracy; rather, its effectiveness depends on the structural properties of the network and characteristics of the recognition objects. In this study, we selected the Acmix attention mechanism due to its superior performance. Acmix possesses the unique capability of dynamically adjusting channel weights by calculating the global attention map for each channel. By contrast, ECA (Wang et al., 2023), CBAM (Woo et al., 2018), and SE (Hu et al., 2018) employ fixed channel weights. This dynamic adjustment allows Acmix to highlight critical channel information more precisely, effectively enhancing the features of the walnut images. Additionally, Acmix can concurrently capture a more comprehensive spatial-frequency feature representation by combining location and channel attention information. The introduction of location attention further promotes channel attention, thereby enhancing the extraction of pertinent location features from walnut images. Conversely, ECA, CBAM, and SE consider only a single type of attention, whether it is location, position, or channel. Moreover, although model size and inference speed are crucial considerations, accuracy remains paramount. As depicted in Table 6, the additional model burden introduced by Acmix was a mere 3.2 MB, and the reduction in FPS was a modest 4.08 frames per second. The increase in the parameters, although present, does not overly complicate the model. Given that the Acmix hybrid mechanism comprehensively captures information, combining both spatial and channel contextual insights and significantly enhancing mAP, the slight reduction in computational efficiency and detection speed remains acceptable. Although the task of simultaneously increasing detection accuracy while maintaining model efficiency is inherently challenging, the experimental findings suggest that

TABLE 6 Comparison of target detection model capabilities with the addition of various focus mechanisms.

Attention mechanisms	mAP0.5 (%)	Parameters	Model size (MB)	FPS
None	78.97	7,276,605	27.1	47.22
ECA	79.15	7,276,866	27.1	50.95
CBAM	78.90	7,278,751	27.1	51.88
SE	79.69	7,277,629	27.1	51.03
Acmix	80.35	8,106,537	30.3	43.14

incorporating attention mechanisms can mitigate this challenge to some extent. For instance, Table 6 illustrates that the inclusion of attention modules (e.g., Acmix, ECA, etc.) can effectively improve the mAP with minimal expansion in model size. Among these mechanisms, Acmix attention stands out by achieving the best accuracy improvement, driven by its ability to integrate spatial and channel contextual information. Considering all factors related to model accuracy, size, and detection speed, the Acmix attention mechanism emerges as the optimal choice, striking an excellent balance between accuracy enhancement, model size, and detection speed.

3.5 Enhancing detection performance for varied target sizes

The classification of good and bad walnuts was notably affected by the degreening and drying process. Walnuts typically sold fall within the size range of 20–50 mm, and the initial processing of walnuts after degreening and drying significantly influences their classification. It is particularly crucial to ensure a sufficiently large field of view in the context of multi-target walnuts to improve grading efficiency. Additionally, evaluating the recognition performance of the model for multi-target walnuts in an actual mixed forward conveying scenario represents a rigorous test of its capabilities. To investigate the performance of the detection model for multi-target walnuts of varying sizes in a mixed scene, as illustrated in Figure 9, we employed 30 small target walnuts measuring 20–30 mm, 30 medium target walnuts ranging from 30–40 mm, and 30 large target walnuts spanning 40–50 mm. Each size category included 10 good walnuts and 20 bad walnuts (10 with black spots and 10 broken fruits). Additionally, we incorporated 30 walnuts ranging in size from 20 to 50 mm (10 walnuts per size), featuring 3 good walnuts and 7 bad walnuts in each size (4 with black spots and 3 broken walnuts). Comparing the small-target detection results in Figures 9E, I, it becomes evident that the improved model identified good walnuts within the small-target category. This improvement can be attributed to the replaced MobileNetV3 module, which effectively captures multiscale information through depth-separable convolution, enhancing the recognition of key features, such as the morphology of small target walnuts. Upon comparing the target images in Figures 9F, J, the large target images in Figures 9G, K, and the mixed target detection images in Figures 9H, L, it becomes apparent that the original model struggled to identify good walnuts, particularly in the case of multiple targets, in which simultaneous identification was problematic. By contrast, the improved model adeptly identified each individual walnut, significantly enhancing the detection accuracy of healthy fruits. This improvement can be attributed to the addition of the Acmix attention mechanism after the SPP layer, which effectively captures spatial feature information related to walnut fruit shape and surface texture across multiple scales. Meanwhile, the MetaAcon activation function is more effective at expressing non-linear features than SiLU, enabling the extraction of complex features, such as walnut fruit color, and aids in the identification of walnuts of varying sizes. Therefore, the improved

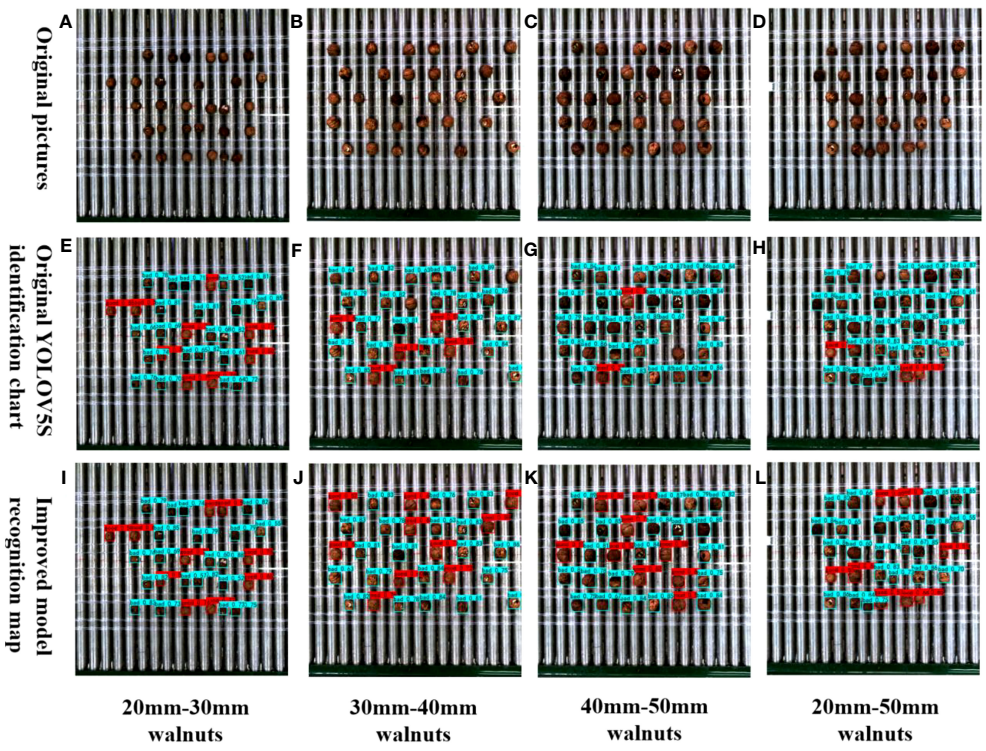


FIGURE 9
Comparison between the original and improved models for detecting various sizes of mixed walnuts.

YOLOv5s_AMM model demonstrates enhanced recognition performance in the mixed recognition of multi-target walnuts at different scales, maintaining a high recognition count and rate. Although a single false detection occurred in the small-target scene, the overall false detection rate remained below 3%. Future research efforts will address this issue by refining the structure of the model to detect small-sized targets.

3.6 Ablation experiments

Ablation experiments were performed on the original YOLOv5s model to assess the impact of various enhancement strategies on the detection performance. All improvement procedures were trained and validated using identical training and validation datasets, and evaluated using the same test dataset. The experimental results are presented in Table 7. The original YOLOv5s model achieved an mAP of 78.97% based on 201 test images. It featured a parameter count of 7,276,605, model size of 27.1 MB, and FPS of 47.22 frames per second, as detailed in Table 7. Notably, the three enhancements proposed in this study yielded positive effects on multiple facets of the original model. First, replacement of the C3 structure in MobileNetV3's backbone network resulted in a notable mAP increase of 1.54%, reaching 80.51%. This enhancement also significantly reduced the footprint of the model by 37.6%, decreased the number of parameters by 39.6%, and boosted the frame rate by 3.42 frames per second. Subsequently, the addition of the Acmix module further improved the mAP and frame rate of the model, albeit not to the same extent as the inclusion of MobileNetV3 alone. Finally, the integration of the MetaAconc module into the model facilitated an adaptability to specific input data during training, culminating in an enhanced performance across various tasks. This configuration achieved an mAP of 80.78%, featuring 5,424,971 parameters, a model size of 20.9 MB, and a frame rate of 40.42 frames per second. Following network optimization, the mAP experienced a 0.59% improvement, with no changes in the total number of model parameters and detection speed. Compared with the original YOLOv5 model, the enhanced model demonstrated a 1.81% improvement in mAP, a substantial 22.88% reduction in model size, and a notable 25.45% decrease in the number of parameters. In conclusion, the method proposed in this study delivers a rapid high-accuracy detection performance across small and large scales. This meets the requirements for real-time detection while maintaining a compact model size.

3.7 Comparative experiments

In this study, we retrained several conventional network models to assess the performance differences between the improved models and their established counterparts. We employed a control-variable approach to ensure the accuracy of the computational results. Subsequently, we compared the detection results of the various network models using the same test dataset. The comparative results are presented in Table 8, highlighting the disparities in the mAP detection performance, model size, and detection speed. For multi-target kernel detection, our improved model achieved the highest recognition accuracy, surpassing the original YOLOv5s model. Specifically, it outperformed the YOLOv4_tiny (75.47%), EfficientNet_YOLOv3 (75.95%), MobileNetV1_YOLOv4 (73.77%), YOLOv3 (80.56%), and YOLOv4 (80.52%) models by 1.81%, 5.31%, 4.83%, 7.01%, 0.22%, and 0.26%, respectively. Concerning parameter count, our improved model stood out with only 5,424,971 parameters, which was significantly lower than the other comparison models. In terms of model size, our model's footprint was merely 20.9 MB, making it the most compact, in stark contrast to the YOLOv4 model's size of 244 MB and the YOLOv3 model's size of 235 MB. Furthermore, our improved model achieved a detection frame rate of 40.42 frames per second, surpassing the YOLOv4 model by 7.77 frames per second and EfficientNet_YOLOV3 by 8.32 frames per second. In summary, our enhanced lightweight walnut detection model excels in recognition accuracy, boasts a compact model size, and demonstrates superior inference speed compared with its counterparts.

As shown in Figure 10, We selected images captured from the actual primary processing grading equipment to represent different sizes of walnuts, including small targets (20–30 mm), medium targets (30–40 mm), Each size category comprised 10 good and 20 bad fruits, with 10 each of black spots and broken fruits. Additionally, we included 30 walnuts ranging from 20 to 50 mm

TABLE 7 Impact of various enhancement strategies on model performance.

Model	mAP@0.5 (%)	Parameters	Model size (MB)	FPS
YOLOv5s	78.97	7,276,605	27.1	47.22
+Acmix	80.35	8,106,537	30.3	43.14
+MobileNetV3	80.51	4,395,327	16.9	50.64
+MetaAconc	80.18	47,098,541	27.2	49.64
+Acmix+MobileNetV3	80.19	5,412,267	20.8	39.34
+Acmix+MetaAconc	79.90	7,908,875	30.3	42.33
+MobileNetV3+MetaAconc	79.91	4,595,039	17.7	52.26
+Acmix+MobileNetV3+MetaAconc	80.78	5,424,971	20.9	40.42

TABLE 8 Detection results of various target detection algorithms on walnut images.

Model	mAP@0.5 (%)	Parameters	Model size (MB)	FPS
YOLOv5s_Acmix_ MobileNetV3_MetaAconc (ours)	**80.78**	**5,424,971**	**20.9**	40.42
YOLOv5s	78.97	7,276,605	27.1	47.22
YOLOv4_tiny	75.47	6,056,606	22.4	**111.48**
EfficientNet_YOLOv3	75.95	10,776,233	60	32.1
MobileNetV1_YOLOv4	73.77	12,692,029	51.1	57.06
YOLOv3	80.56	61,949,149	235	48.23
YOLOv4	80.52	64,363,101	244	32.65

The bold values indicate the optimal values corresponding to the four groups of data: map@0.5(%)80.78 has the highest precision and is marked with bold; The number of parameters is 5,424,971, with the minimum marking thickness; Model size 20.9MB, minimum size; FPS11.48, the fastest detection speed.

(comprising large, medium, and small sizes of 10 each), featuring 3 good and 7 bad fruits (4 with black spots and 3 broken fruits) for each size. Subsequently, we compared and examined the true results for each walnut size and category (Table 9). The experimental findings, in terms of the identification of good, bad, unchecked, and incorrectly detected walnuts, affirm the improved YOLOv5s_AMM model's efficacy and precision in discerning good and bad walnuts across large, medium, and small targets. Remarkably, there were minimal instances of unchecked and incorrectly detected walnuts of different sizes. Notably, the detection of small target walnuts, characterized by a complex surface morphology and small size, poses a significant challenge. Although the YOLOv5s, YOLOv4_tiny, and YOLOv4 models exhibited relatively similar recognition results to the improved model, occasional cases of non-detection and incorrect detection were observed, underscoring the improved model's superiority. Comparatively, the YOLOv4_tiny and EfficientNet_YOLOv3 models displayed slightly better results than the improved model, but with a notable increase in false detections and non-detections. Conversely, models such as YOLOv3 and the original YOLOv5s demonstrated ineffectiveness at detecting good fruits, with a significant number of non-detections and false detections. In conclusion, the enhanced YOLOv5s_AMM model consistently demonstrated its effectiveness and precision in identifying good and bad walnut fruits across varying sizes, as assessed by a composite set of criteria encompassing good and bad fruit identification and unchecked and incorrectly detected walnuts.

4 Discussion

This study introduces a rapid non-destructive detection model designed to enhance the performance of the original YOLOv5s network for the detection of good and bad fruits within multi-target samples of dried walnuts. The dataset encompassed specimens that exhibited both desirable attributes and imperfections, including instances of black spotting and breakage. It is worth noting that extant research concerning walnuts predominantly centers on fresh green walnuts or kernels obtained from orchard trees, as shown in Table 10. Historical limitations have constrained access to extensive repositories of high-resolution imagery depicting good and bad dried walnuts, thereby constraining the scope of deep learning investigations in this domain. Recently, there has been a prevalent shift towards machine learning and convolutional neural networks in the context of kernel defect detection. The datasets used in related literature predominantly consist of single-object images captured within controlled laboratory environments or images featuring multiple object sets against the backdrop of orchard settings. In the interest of equitable assessment pertaining to image composition, network architecture, and detection efficacy within the chosen dataset, the findings presented in Table 10 elucidate discernible enhancements in detection performance achieved through the deployment of various optimized network configurations relative to the original model. Consequently, these results underscore both the effectiveness and necessity of augmenting the detection capabilities of networks tasked with discerning multiple objects of diverse sizes.

In this study, we analyzed the experimental results obtained from the improved YOLOv5s_AMM detection model. The primary focus of this study was to address the challenge of discerning good and bad fruit images among multi-target walnuts of varying sizes. Moreover, we assessed the recognition performance of the improved model across walnut images with different dimensions. Within this analysis, we explored the impact of various attention mechanisms (Table 11) and the influence of different positions of improvement (Table 6) on the model's recognition capabilities. Notably, the enhancements made to the original YOLOv5s model encompassed the incorporation of the Acmix structure, which introduces convolutional mixing, following the SPP layer. In addition, the activation function within the neck layer convolution was replaced with the MetaAconC activation function. These improvements were substantiated by the ablation (Table 7) and comparative experiments (Table 8). The experimental results presented in this study demonstrate the ability of the enhanced YOLOv5s_AMM detection model to swiftly and accurately identify good and bad walnuts within mixed images of dried walnuts, encompassing multiple targets of varying sizes. Furthermore, comparative experiments involving diverse improved modules and different typical target detection networks contribute to a comprehensive evaluation of the proficiency of the model in recognizing good and bad walnut fruits.

10.3389/fpls.2023.1247156

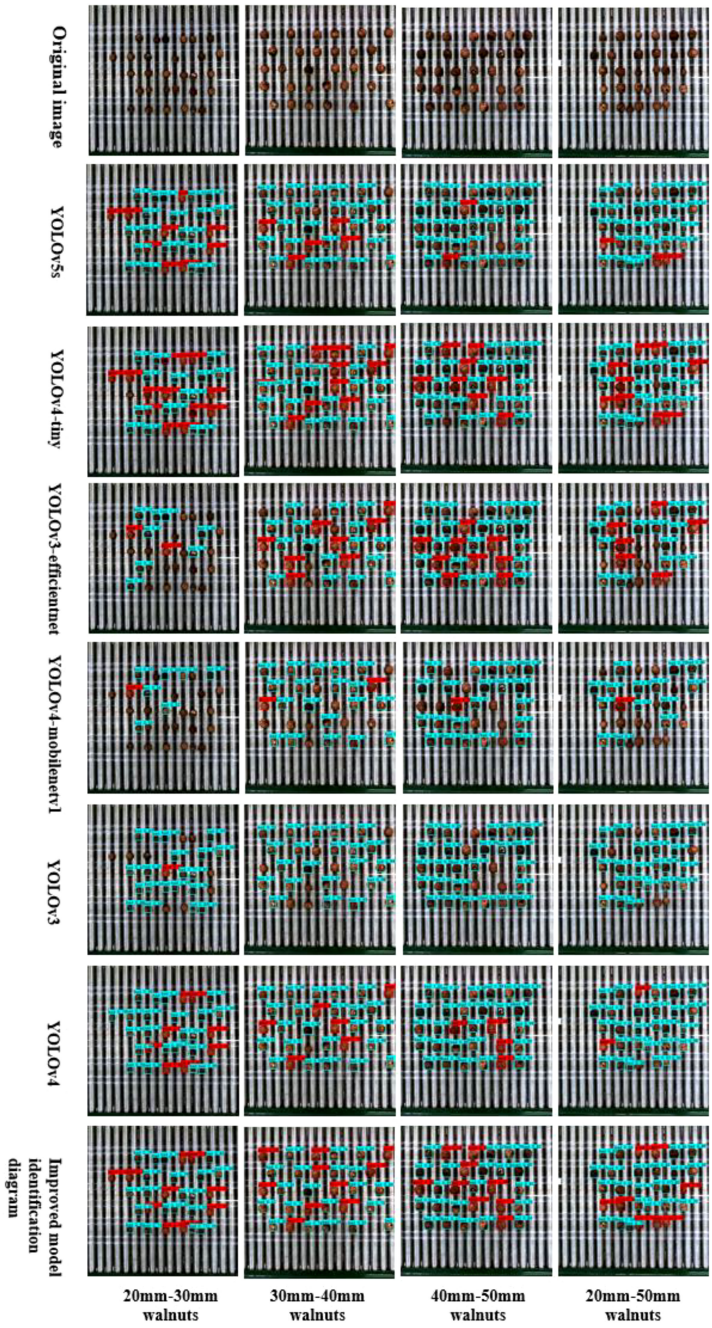


FIGURE 10
Detection performance of mainstream detection networks for good and defective walnuts of various sizes.

In essence, our proposed enhanced network exhibits improved detection performance, reduced model size, and accelerated inference speed when tasked with identifying mixed multi-target dried walnut fruits of varying sizes. This characteristic holds a significant promise for deployment in resource-constrained edge devices. In future research endeavours, we plan to prioritize the refinement of recognition accuracy and the model's generalization capabilities. This will entail extending its applicability to encompass a broader spectrum of walnut variety recognitions. Subsequently, we aim to implement an improved model within the grading equipment used in the primary processing stages of walnuts. This deployment is envisaged to not only augment the value of walnut products but also enhance the efficiency of the walnut industry's grading processes.

5 Conclusions and future work

This study focused on using photographs of walnuts collected after degreening, cleaning, and drying as the research dataset. In response to the distinctive visual attributes of walnuts within the primary processing context, we developed and implemented a

TABLE 9 Number of good and bad walnuts detected by mainstream target detection networks for various walnut sizes.

Model	20mm–30mm good/bad/ uninspected/ misinspected	30mm–40mm good/bad/ uninspected/ misinspected	40mm–50mm good/bad/ uninspected/ misinspected	20mm–50mm good/bad/ uninspected/ misinspected
YOLOv5s_Acmix_MobileNetV3_MetaAconc (ours)	**10/20/0/0**	**10/20/0/0**	**9/21/0/1**	**8/22/0/2**
YOLOv5s	9/21/0/1	5/23/2/3	2/27/1/8	3/26/1/4
YOLOv4_tiny	13/26/1/3	10/20/0/4	8/21/1/1	8/20/2/3
EfficientNet_YOLOv3	2/8/20/0	8/16/6/0	9/20/1/3	6/14/10/3
MobileNetV1_YOLOv4	1/7/22/2	2/19/9/3	1/21/8/1	1/14/15/1
YOLOv3	1/20/9/1	0/22/8/3	0/25/5/0	0/23/7/3
YOLOv4	8/22/0/2	6/23/1/3	4/24/0/5	2/27/1/5

The bold values represent the best case for each of these four sets of cases [best for good, bad, undetected, false (quantity) data in different sizes (20-30\30-40\40-50\20-50 \20-50).

TABLE 10 Recent research on target detection in walnut studies.

Objects	Networks	Dataset conditions	mAP	F1	Accuracy
Walnut foreign body (Rong et al., 2019)	Machine vision combines two different convolutional neural networks	Walnuts, natural foreign objects, and artificial foreign objects	—	—	95%
Pecan abscission, shell, and embryo area (Costa et al., 2021)	Machine vision combined with Mask-RCNN	Abscission, shell, and embryo areas in both small (young) and large (old) pecans at multiple growth stages	—	95.3% ~100%	—
Green walnut in natural environments (Fan et al., 2021)	Improved and faster R-CNN	Detection of green walnuts in natural environments (uniform light, uneven light, overlapping objects, shading, and varying target sizes)	97.71%	96.12%	—
Green walnut in natural environments (Hao et al., 2022)	Improved YOLOv3 (MobileNetV3)	Green walnuts on trees in the orchard (large targets, small targets, and backlighting conditions)	86.11%	—	—
English walnut kernel pericarp colour (Donis-Gonzalez et al., 2020)	Machine vision combined with a stepwise logistic regression method	English walnut kernels with different coloured peels	—	—	"Chandler" model (88.8%), seedling model (80.4%), and "Howard" model (75.1%)
Walnut impurities (Yu et al., 2023)	Improved YOLOv5 (Transformer and GhostNet)	Small impurities within walnut kernels	88.9%	—	—
Walnut fresh fruit (Zhang et al., 2016)	Machine vision combining hybrid features with the least squares support vector machines	Identification of fresh pecan fruits under natural scenes, considering downlight backlighting and branch shading	—	—	92.48%

TABLE 11 Comparison of the effects of adding the attention mechanism at different positions.

Applied position	mAP0.5 (%)	Parameters	Model size (MB)	FPS
None	78.97	7,276,605	27.1	47.22
Position 1	80.35	8,106,537	30.3	43.14
Position 2	78.90	7,833,645	29.2	28.39
All	79.69	8,663,577	32.4	26.57

method for detecting multiple good and bad walnut fruit targets. To support this investigation, we collected a substantial volume of multi-target walnut images, thereby constructing a corresponding dataset. To enhance the efficiency of the model while maintaining its lightweight architecture, we replaced the C3 network in the original YOLOv5s with MobileNetV3, resulting in an M3-Net network. Subsequently, we explored the impact of various attention mechanisms and improvement positions on the walnut images. Notably, the Acmix structure after the SPP layer was introduced, integrating attention and hybrid modules to capture both global and local image features. This strategic modification reduces network computational costs while augmenting performance on complex datasets. Furthermore, the MetaAconC activation function of the CBM module in the neck layer was replaced with an SiLU activation function from the original Conv layer. This adaptation improved the distribution of input-specific data for fine-tuning, thereby enhancing feature detection across various image scales. Additionally, we assessed the effectiveness of the model across the walnut images in varying proportions. Finally, we conducted a comprehensive examination of the different improvement modules applied to the detection of walnut datasets within the backbone and neck layers of the Ai model. The performance of different target detection networks on walnut datasets were further investigated. The results of these experiments successfully validated the performance enhancements achieved by our improved model.

The principal findings of this study are summarized as follows:

(1) Compared with other target detection models, our improved model exhibited superior performance across multiple metrics, including detection precision, model size, parameter size, and detection speed. Notably, our improved model achieved the highest accuracy, with an mAP of 80.78. Moreover, it boasted the smallest model size, measuring 20.9 MB, which was notably 11.7 times and 11.2 times smaller than the model sizes of conventional algorithms such as YOLOv4 and YOLOv3, respectively. Simultaneously, the model maintained a detection speed of 40.42 frames per second, aligning with the lightweight nature of the model suitable for rapid walnut detection scenarios and substantially outperforming the YOLOv4 and YOLOv3 models in terms of speed. These results underscore the success of the improved model in achieving greater recognition accuracy, a compact model size, and rapid performance.

(2) In practical applications, the enhanced model was employed to distinguish between good and bad fruits of multi-target walnuts within the test set. Ablation experiments were conducted to assess its performance, which resulted in an mAP of 80.78%. Compared with the original YOLOv5s model, our enhanced model exhibited an increase of 1.81% in mAP, a reduction of 22.88% in model size, and a decrease of 25.45% in parameter count, while maintaining a negligible difference in FPS. Additionally, experimental results involving walnut image detection with varying target sizes indicate improved precision and robustness.

(3) By leveraging the capabilities of the improved YOLOv5s_AMM model, which addresses the gap in detecting walnuts of different sizes after peeling and drying in the preliminary processing stage, we intend to apply it to the preliminary processing operations of walnut processing enterprises. Specifically, the model was employed for the detection and grading of good and bad walnut fruits after the peeling, washing, and drying stages. Our model offers distinct advantages, including a high recognition accuracy and compact model size.

Data availability statement

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

Author contributions

ZZ designed a lightweight model and trained the model. HZ walnut multi-target image acquisition; LL and YL lead many experiments and revisions to the text; ZL and HZ labeled walnut data set; XL to help build collection equipment and purchase materials; FZ and YZ received guidance and financial support for their experiments. All authors contributed to this article and approved the submitted version.

Funding

This research is supported by Yunnan Forestry and Grassland Comprehensive Technology Extension Station (2022530101000350), Kunming University of Science and Technology 2022 Graduate Talents Enhancement Program (CA22369M0A) and Yunnan Academician (Expert) Workstation Financial support of Project (202105AF150030).

Acknowledgments

We are extremely appreciative of all the authors' support and contributions to the manuscript.

Conflict of interest

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

Publisher's note

References

Arunkumar, M., Rajendran, A., Gunasri, S., Kowsalya, M., and Krithika, C. (2021). Non-destructive fruit maturity detection methodology-A review. *Materials Today: Proc.* doi: 10.1016/j.matpr.2020.12.1094

Bhattacharya, U., Ghassemlooy, Z. J. F. O. C., and Sciences, D. (2006). Overview of ShuffleNet and its variations in lightwave networks. *Foundations of Computing & Decision Sciences* 31, 193–220.

Bochkovskiy, A., Wang, C. Y., and Liao, H. Y. M. (2020). YOLOv4: optimal speed and accuracy of object detection. *arXiv preprint arXiv:2004.10934*. doi: 10.48550/arXiv.2004.10934

Chakraborty, S. K., Subeesh, A., Dubey, K., Jat, D., Chandel, N. S., Potdar, R., et al. (2023). Development of an optimally designed real-time automatic citrus fruit grading–sorting machine leveraging computer vision-based adaptive deep learning model. *Eng. Appl. Artif. Intell.* 120, 105826. doi: 10.1016/j.engappai.2023.105826

Cobus, L. A., and van Wijk, K. (2023). Non-contact acoustic method to measure depth-dependent elastic properties of a kiwifruit. *Wave Motion* 119, 103126. doi: 10.1016/j.wavemoti.2023.103126

Costa, L., Ampatzidis, Y., Rohla, C., Maness, N., Cheary, B., and Zhang, L. (2021). Measuring pecan nut growth utilizing machine vision and deep learning for the better understanding of the fruit growth curve. *Comput. Electron. Agric.* 181, 105964. doi: 10.1016/j.compag.2020.105964

Donis-Gonzalez, I. R., Bergman, S. M., Sideli, G. M., Slaughter, D. C., and Crisosto, C. H. (2020). Color vision system to assess English walnut (Juglans Regia) kernel pellicle color. *Postharvest Biol. Technol.* 167, 111199. doi: 10.1016/j.postharvbio.2020.111199

Fan, X., Xu, Y., Zhou, J., Liu, X., and Tang, J. I. (2021). Green walnut detection method based on improved convolutional neural network. *Trans. Chin. Soc. Agric. Machinery* 52, 149–155+114. doi: 10.6041/j.issn.1000-1298.2021.09.017. Y. W.

Hao, J., Bing, Z., Yang, S., and Yang, J. (2022). Detection of green walnut by improved YOLOv3. *Trans. Chin. Soc. Agric. Eng.* 38, 183–190. doi: 10.11975/j.issn.1002-6819.2022.14.021. L. S.

He, K., Zhang, X., Ren, S., and Sun, J. (2015). Spatial pyramid pooling in deep convolutional networks for visual recognition. *IEEE transactions on pattern analysis and machine intelligence* 37, 1904–1916. doi: 10.1109/TPAMI.2015.2389824

Howard, A., Ruoming, P., Adam, H., Quoc, L., Sandler, M., Bo, C., et al. (2019). "Searching for mobileNetV3," in *2019 IEEE/CVF International Conference on Computer Vision (ICCV). Proceedings*. 1314–1324. doi: 10.1109/iccv.2019.00140

Hu, J., Shen, L., and Sun, G. (2018). "Squeeze-and-excitation networks," in *2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*. 7132–7141. doi: 10.48550/arXiv.1709.01507

Li, C., Zhang, S., Sun, H., Chen, C., Xing, S., and Zhao, X. (2019). Walnut appearance defect detection based on computer vision. *Modern Food Sci. Technol.* 35, 247–253 +246. doi: 10.13982/j.mfst.1673-9078.2019.8.035

Ma, N. N., Zhang, X. Y., Liu, M., Sun, J.IEEE COMP, S. O. C (2021). "Activate or not: learning customized activation," in *IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, Jun 19-25 2021. 8032–8042, Electr Network. doi: 10.1109/cvpr46437.2021.00794

Nan, Y. L., Zhang, H. C., Zeng, Y., Zheng, J. Q., and Ge, Y. F. (2023). Intelligent detection of Multi-Class pitaya fruits in target picking row based on WGB-YOLO network. *Comput. Electron. Agric.* 208, 107780. doi: 10.1016/j.compag.2023.107780

Pan, X. R., Ge, C. J., Lu, R., Song, S. J., Chen, G. F., Huang, Z. Y., et al. (2022). "On the integration of self-attention and convolution," in *IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, New Orleans, LA, Jun 18-24 2022. 815–825. doi: 10.1109/cvpr52688.2022.00089

Rong, D., Xie, L. J., and Ying, Y. B. (2019). Computer vision detection of foreign objects in walnuts using deep learning. *Comput. Electron. Agric.* 162, 1001–1010. doi: 10.1016/j.compag.2019.05.019

Sandler, M., Howard, A., Zhu, M. L., Zhmoginov, A., Chen, L. C.IEEE (2018). "MobileNetV2: inverted residuals and linear bottlenecks," in *31st IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, Salt Lake City, UT, Jun 18-23 2018. 4510–4520. doi: 10.1109/cvpr.2018.00474

Su, L., Xie, Z., Xiong, X., Liang, L., Qi, H., Yang, Y., et al. (2021). Design principle and process of yunnan green walnut peeling machine. *Guangxi Forestry Sci.* 50, 218–222. doi: 10.19692/j.cnki.gfs.2021.02.017

Sustika, R., Subekti, A., Pardede, H. F., Suryawati, E., Mahendra, O., and Yuwana, S. (2018). Evaluation of deep convolutional neural network architectures for strawberry quality inspection. *Int. J. Eng. Technol.* 7, 75–80. doi: 10.14419/ijet.v7i4.40.24080

Taylor, L., and Nitschke, G. (2018). "Improving deep learning with generic data augmentation," in *8th IEEE Symposium Series on Computational Intelligence (IEEE SSCI)*, Bengaluru, INDIA, Nov 18-21 2018. 1542–1547. doi: 10.1109/SSCI.2018.8628742

Wang, D. D., Dai, D., Zheng, J., Li, L. H., Kang, H. Y., and Zheng, X. Y. (2023). WT-YOLOM: an improved target detection model based on YOLOv4 for endogenous impurity in walnuts. *Agronomy-Basel* 13, 1462. doi: 10.3390/agronomy13061462

Woo, S. H., Park, J., Lee, J. Y., and Kweon, I. S. (2018). "CBAM: convolutional block attention module," in *15th European Conference on Computer Vision (ECCV)*, Munich, GERMANY, Sep 08-14 2018. 3–19. doi: 10.1007/978-3-030-01234-2_1

Xie, L., Li, X., Ning, D., Ma, T., and Wu, T. (2021). Research progress on germplasm resources evaluation and excavation of deep striated walnut. *For. Sci. Technol.* 6, 26–31. doi: 10.13456/j.cnki.lykt.2021.02.22.0001

Yu, L., Qian, M., Chen, Q., Sun, F., and Pan, J. (2023). An improved YOLOv5 model: application to mixed impurities detection for walnut kernels. *Foods* 12, 624. doi: 10.3390/foods12030624

Zhang, Y., Zhou, Z., and Zhou, S. (2016). Segmentation method for fresh hickory nut recognition based on mixed features and LS-SVM classifier. *J. Cent. South Univ. Forestry Technol.* 36, 137–143. doi: 10.14067/j.cnki.1673-923x.2016.11.024

frontiers | Frontiers in Plant Science

TYPE Original Research
PUBLISHED 16 November 2023
DOI 10.3389/fpls.2023.1324152

OPEN ACCESS

EDITED BY
Lie Deng,
Southwest University, China

REVIEWED BY
Yi Yang,
Beijing Technology and Business University, China
Leiqing Pan,
Nanjing Agricultural University, China

*CORRESPONDENCE
Huijie Peng
779642214@qq.com
Jiangbo Li
lijb@nercita.org.cn

RECEIVED 19 October 2023
ACCEPTED 30 October 2023
PUBLISHED 16 November 2023

CITATION
Mei M, Cai Z, Zhang X, Sun C, Zhang J, Peng H, Li J, Shi R and Zhang W (2023) Early bruising detection of 'Korla' pears by low-cost visible-LED structured-illumination reflectance imaging and feature-based classification models. *Front. Plant Sci.* 14:1324152. doi: 10.3389/fpls.2023.1324152

Early bruising detection of 'Korla' pears by low-cost visible-LED structured-illumination reflectance imaging and feature-based classification models

Mengwen Mei[1], Zhonglei Cai[1], Xinran Zhang[1], Chanjun Sun[2], Junyi Zhang[1], Huijie Peng[1,3,4]*, Jiangbo Li[5,6]*, Ruiyao Shi[5] and Wei Zhang[7]

[1]College of Mechanical and Electrical Engineering, Shihezi University, Shihezi, China, [2]Jiangsu Province and Education Ministry Co-sponsored Synergistic Innovation Center of Modern Agricultural Equipment, Jiangsu University, Zhenjiang, China, [3]Xinjiang Production and Construction Corps Key Laboratory of Modern Agricultural Machinery, Shihezi, China, [4]Engineering Research Center for Production Mechanization of OasisCharacteristic Cash Crop, Ministry of Education, Shihezi, China, [5]Intelligent Equipment Research Center, Beijing Academy of Agriculture and Forestry Sciences, Beijing, China, [6]National Engineering Research Center for Information Technology in Agriculture, Beijing, China, [7]Department of Computer Technology and Science, Anhui University of Finance and Economics, Bengbu, China

Introduction: Nondestructive detection of thin-skinned fruit bruising is one of the main challenges in the automated grading of post-harvest fruit. The structured-illumination reflectance imaging (SIRI) is an emerging optical technique with the potential for detection of bruises.

Methods: This study presented the pioneering application of low-cost visible-LED SIRI for detecting early subcutaneous bruises in 'Korla' pears. Three types of bruising degrees (mild, moderate and severe) and ten sets of spatial frequencies (50, 100, 150, 200, 250, 300, 350, 400, 450 and 500 cycles m^{-1}) were analyzed. By evaluation of contrast index (CI) values, 150 cycles m^{-1} was determined as the optimal spatial frequency. The sinusoidal pattern images were demodulated to get the DC, AC, and RT images without any stripe information. Based on AC and RT images, texture features were extracted and the LS-SVM, PLS-DA and KNN classification models combined the optimized features were developed for the detection of 'Korla' pears with varying degrees of bruising.

Results and discussion: It was found that RT images consistently outperformed AC images regardless of type of model, and LS-SVM model exhibited the highest detection accuracy and stability. Across mild, moderate, severe and mixed bruises, the LS-SVM model with RT images achieved classification accuracies of 98.6%, 98.9%, 98.5%, and 98.8%, respectively. This study showed that visible-LED SIRI technique could effectively detect early bruising of 'Korla' pears, providing a valuable reference for using low-cost visible LED SIRI to detect fruit damage.

KEYWORDS

pears, early bruise detection, classification, machine learning, visible LED structured illumination

1 Introduction

Bruising is the most common type of mechanical damage (Opara and Pathare, 2014), particularly on fruit like pears which are sensitive to mechanical damage (Celik, 2017). Bruises may occur when the stress on the fruit surface exceeds the failure stress of fruit tissue. It is a kind of subcutaneous tissue injury without rupture of the skin of fruit (Opara and Pathare, 2014; Hussein et al., 2019; Mei and Li, 2023). The formation of bruise will not only lead to physiological changes in fruit density, moisture content, browning degree and firmness, but also accelerate the respiration rate of fruit and increase the production of ethylene (Khurnpoon and Siriphanich, 2011; Polat et al., 2012; Bian et al., 2020), thereby accelerating the decay process of fruit and leading to significant economic losses. However, the non-destructive and accurate detection of early bruised fruit is extremely challenging.

Many techniques have been used for bruising detection of pears, including hyperspectral imaging (HSI) (Lee et al., 2014; Fu and Wang, 2022; Tian et al., 2023), magnetic resonance imaging (MRI) (Razavi et al., 2018; Razavi et al., 2020), X-ray computed tomography (CT) (Azadbakht et al., 2019a; Azadbakht et al., 2019b), thermal imaging (TI) (Kim et al., 2014; Zeng et al., 2020), Optical coherence tomography (OCT) (Zhou et al., 2019), etc. HSI has been widely used in fruit damage detection and has been proven effective in this regard. However, its capabilities for detection of early-stage bruises, especially immediate post-bruise detection, still require enhancement. Additionally, HSI is too slow and expensive for commercial applications (Tian et al., 2021). For MRI, CT, OCT, they can capture high contrast images but equipment cost is an important consideration factor in practical applications. TI is a detection technology that does not require a light source. It can non-invasively convert the radiation of an object into a surface temperature distribution for bruising detection (Zeng et al., 2020). However, it has strict temperature requirements, and the fruit may be affected by the heating/cooling process.

Traditional imaging systems (e.g. HSI, multispectral imaging and machine vision) commonly used uniform or diffuse illumination for fruit quality detection, making it difficult to control light penetration and interaction with biological tissue, which limits their performance in detecting depth-specific information such as subsurface tissue bruising in fruit (Lu and Lu, 2017; Lu and Lu, 2019). Structured illumination (SI) can be used to enhance the detection of subsurface defects in fruit by varying the spatial frequency of the illumination to control the depth of light penetration into the tissue (Li et al., 2023; Li et al., 2024). Depending on the purpose of the application, SI techniques can be implemented using either inverse or forward methods. Spatial frequency domain imaging (SFDI) based on inverse methods can be used to obtain absorption coefficients and approximate scattering coefficients of fruit tissue by means of inverse algorithmic diffusion models (Sun et al., 2019). This method has also been used for bruise detection in pears (He et al., 2018). Different from SFDI, structured-illumination reflection imaging (SIRI) is used to enhance the detection of subsurface damage of fruit in a simpler and faster way. The pattern image obtained by demodulation can obtain direct component (DC) and alternating component (AC) images, in which the AC image carries depth resolution information and can be used for the detection of subsurface tissue bruising in fruit (Lu et al., 2016a; Li et al., 2023). SIRI has now been used to detect bruises on apples (Lu et al., 2016a) and pickling cucumbers (Lu et al., 2021) with good results. Recently, our laboratory developed a new SIRI system based on light-emitting diode (LED) light source and monochromatic camera, which can realize fruit detection in visible light band, and further reduce the cost of SIRI system while obtaining good subcutaneous damage detection effect. The system has been used to detect the early decay of oranges (Cai et al., 2022).

The aim of this study was to demonstrate the ability of low-cost visible-LED SIRI to detect pear bruising at an early stage. The specific objectives were to: (1) Acquire DC and AC images for 'Korla' pears with three types of bruising degrees at ten sets of spatial frequencies using a visible-LED SIRI system to determine the optimal spatial frequency combined with a three-phase image demodulation scheme and contrast index analysis; (2) Extract the texture features of AC and ratio (RT) images through the gray level co-occurrence matrix (GLCM) and select the appropriate features based on the random frog algorithm; (3) Develop the least squares support vector machine (LS-SVM), partial least squares discriminant analysis (PLS-DA), and K-nearest neighbor (KNN) classification models combined with selected texture features to classify sound and bruised pears; and (4) Evaluate the independent bruising degree prediction models and mixed bruising degree prediction model to determine the optimal one for classification of bruised 'Korla' pears.

2 Materials and methods

2.1 Sample preparation

'Korla' pears were used in the study. 'Korla' pear is a characteristic fruit in Xinjiang, China. It is famous for its fine flesh, juicy juice and strong aroma. However, the peel of this pear is very thin and easily damaged. The 'Korla' pears were purchased from a local fruit store in Beijing, China. The 'Korla' pear can be roughly divided into two distinct maturation periods, namely, the green maturation period and the yellow maturation period. During the green maturation period, the skin of the pear appears green, while in the yellow maturation period, it turns fully yellow. Over the course of storage, the color of the pear peel undergoes a transition from green to yellow. Most of the pears sold in the fruit store are in the green maturation period, but according to the different sales time, the epidermis of some pears will gradually become yellow, even full yellow, and some pears also may be reddish in color. In this study, the color of the pears was not taken into account during the purchasing process. For all pear samples, green samples accounted for the majority, with a small amount of red or yellow samples. By a simple visual inspection, 403 pears (three pears were used for spatial frequency selection) without external defects were selected as experimental samples. The pear size varies among them. To replicate the real detection environment, this experiment deliberately refrains from making any distinctions.

The static load range of pear fruit during harvesting, storage and transportation is 60-200 newtons (N) (Wu et al., 2013). Therefore, this study selected 50 N, 100 N, and 150 N as the static load pressure level to induce early bruising in pears. Four hundred pears were randomly divided into 4 groups, with 100 in each group, which were sound group (recorded as S0), mild bruise group (recorded as S1), moderate bruise group (recorded as S2) and severe bruise group (recorded as S3). The pears were balanced at room temperature (temperature 24°C, humidity 42%) for 24 h. After that, the 100 pears in the S0 group were not treated. The pears of S1, S2 and S3 were bruised by pressing presses. The pressure probe end of the press is a cylindrical plastic with a diameter of 3 cm and is connected to a pressure sensor with a display screen. Due to the high curvature radius of the equatorial part of 'Korla' pear, it is more vulnerable to form bruises during sorting and packaging. Consequently, the equatorial section of the pear is chosen and subjected to pressure using a press to induce a static pressure bruise. During sample preparation, the pear sample was placed horizontally under the pressure probe of the press. The pear was placed on a sponge-buffered fruit tray and the handle was slowly pressed. When the pressure sensor display reached a specific reading (50N, 100N and 150N represent S1, S2 and S3, respectively), the pressing was stopped. After standing still for 3 seconds, the handle was slowly loosened and the sample was taken out. Figure 1A depicts the RGB images of pears exhibiting three distinct bruise degrees (S1, S2, S3), which also includes the control group (S0). Figure 1B shows the preparation of bruise samples.

2.2 SIRI system and image acquisition

The SIRI system used in this experiment is mainly composed of a digital projector (DLP4500, Texas Instruments, Dallas, TX, United States) with visible LED lights, a monochromatic camera (MV-CA050-10GM, Hangzhou Hikrobot Intelligent Technology Co., Ltd., Hangzhou, China) with an adjustable focal length lens (MVL-MF1628M-8MP, Hangzhou Hikrobot Intelligent Technology Co., Ltd., Hangzhou, China), two polarizers (PL-D50, RAYAN Technology Co., Ltd., Changchun, China), a long-wave pass filter (the cut-off frequency is 450 nm) (GCC-300201, Daheng New Epoch Technology Inc., Beijing, China), an adjustable sample stage (600LW-WT, Shanghai Weimu Automation Equipment Co., Ltd., Shanghai, China) and a computer that can perform sampling and data processing (Cai et al., 2022). The projector and the camera are located directly above the sample to be tested, perpendicular to the horizontal axis. Additionally, a pair of linear polarizers is placed in front of the projector and the camera to eliminate specular reflection. The projector and the camera are connected to the computer through the data line and controlled by the computer through the binding software. The basic composition and real system of the SIRI system is shown in Figure 2. The SIRI system based on LED light and monochrome camera can obtain SI images in the visible light band, which further reduces the equipment cost.

Images were collected immediately after static pressure was applied to each pear. The sample is positioned on the imaging stage with the bruising area facing upward toward the projector and camera. The height of the platform is adjusted to accommodate all sizes of pears before imaging each sample. The distances from the pear sample to the projector and camera was set at approximately 30 cm. Three phase-shifted sinusoidal patterns (with phase offsets of $-2\pi/3$, 0 and $2\pi/3$) in 8-bit bmp format were created in Matlab (The Mathworks, Inc., Natick, MA, USA) and uploaded to the projector control software on the computer, and then imported into the projector for sample illumination. The camera is set to an exposure time of 50 ms to obtain an 8-bit grayscale image for each pattern projected onto the sample.

2.3 Image demodulation and processing

The pear pattern image collected from the SIRI system cannot be directly used for bruising detection, but needs further image demodulation processing. The image demodulation method used in this experiment is a three-phase demodulation (TPD) method. It is a commonly used sinusoidal image demodulation scheme. This method requires three images with equal phase steps for image demodulation. Through the SIRI system, three images are obtained at each frequency, and the phase offsets are $-2\pi/3$, 0 and $2\pi/3$, respectively (Schreiber and Bruning, 2007). Typically, a two-dimensional sinusoidal fringe pattern can be represented as follows:

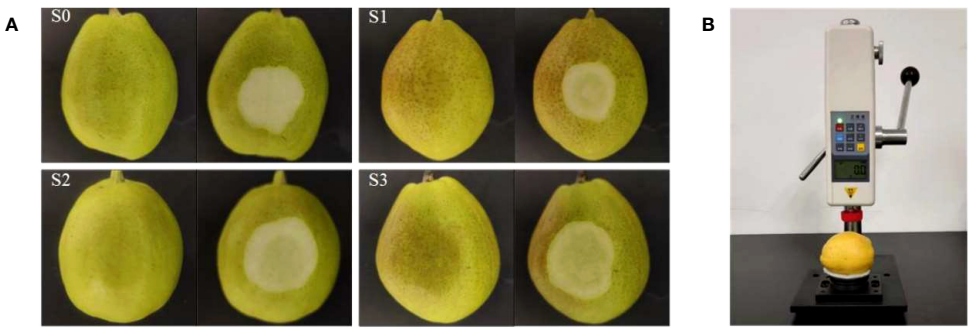


FIGURE 1
(A) Typical 'Korla' pear samples (Unpeeled and peeled) with different degrees of bruising (S0: sound, S1: mild bruises, S2: moderate bruises, S3: severe bruises). **(B)** Preparation of bruise samples.

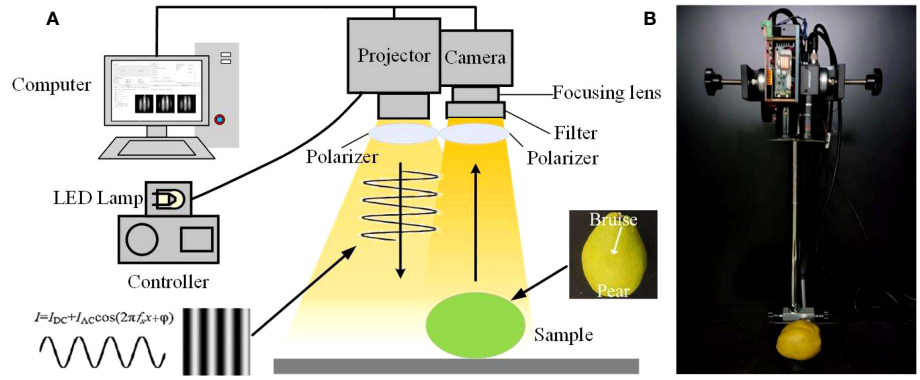


FIGURE 2
Schematic diagram **(A)** and real system **(B)** of the SIRI system.

$$I_n(x,y) = I_{DC}(x,y) + I_{AC}(x,y)\cos(2\pi f_x + 2\pi f_y + \varphi_n) \quad (1)$$

where $I_{DC}(x,y)$ and $I_{AC}(x,y)$ are DC and AC, respectively. f_x and f_y are the spatial frequencies along the x and y axes, respectively. According to the experimental requirements, only the spatial frequency in one direction is required, so f_y is 0 in this experiment. φ_n is the phase shift of the nth pattern image. In this experiment, φ_1, φ_2 and φ_3 corresponding to the sinusoidal fringe patterns $I_1(x,y)$, $I_2(x,y)$ and $I_3(x,y)$ are $-2\pi/3$, 0 and $2\pi/3$, respectively. The DC and AC are the final results obtained by image demodulation, which can be obtained by the following equation (for simplicity, the coordinate symbol is omitted).

$$I_{DC} = \frac{1}{3}(I_1 + I_2 + I_3) \quad (2)$$

$$I_{AC} = \frac{\sqrt{2}}{3}\sqrt{(I_1 - I_2)^2 + (I_1 - I_3)^2 + (I_2 - I_3)^2} \quad (3)$$

The aforementioned equations (3) demonstrate that TPD exclusively relies on straightforward pixel-by-pixel algebraic operations, resulting in efficient computation. Moreover, the subtraction operation effectively mitigates common noise across the three images, enhancing its robustness. The demodulated image DC and AC images correspond to the images acquired under uniform diffuse illumination and the images resulted from the sinusoidal illumination pattern, respectively. The AC image contains depth information, which varies with the spatial frequency of the illumination pattern. Specifically, as the spatial frequency of the illumination patterns increased, the depth of tissue interrogation in the AC images decreased (Lu and Lu, 2019).

Although AC image has the ability of enhanced detection, there are still some deficiencies, such as low intensity, uneven brightness distribution, and large noise. Since DC images also have similar problems, the AC image can be divided by the corresponding DC image to obtain a ratio image RT image to improve the image quality. RT image can make the image background more uniform and enhance the image contrast. It is defined as follows:

$$RT = \frac{I_{AC}}{I_{DC}} = \frac{\sqrt{2}}{I_1+I_2+I_3}\sqrt{(I_1 - I_2)^2 + (I_1 - I_3)^2 + (I_2 - I_3)^2} \quad (4)$$

2.4 Spatial frequency selection

Since the different penetration depths of structured light at different spatial frequencies, it is crucial to select the appropriate frequency for accurate detection of pear bruises. Through preliminary experiments, it was found that the detection effect of bruising was good when the spatial frequency was between 0 and 500 cycles m^{-1}. Therefore, the spatial frequencies of 50, 100, 150, 200, 250, 300, 350, 400, 450 and 500 cycles m^{-1} were selected for imaging, and the optimal frequency that can accurately detect the bruises was selected by comparing the demodulation results. Prior to conducting the experiment, it is essential to generate sinusoidal fringe patterns with varying spatial frequencies on the computer. The generation formula is presented in equation (1). The value I_{DC} and I_{AC} were set to (255/2). By adjusting the parameters f_x or f_y within the equation, fringe patterns corresponding to different spatial frequencies can be generated. These fringe patterns are visually recognizable, appearing as densely-packed black and white stripes at higher spatial frequencies, and sparser black and white stripes at lower spatial frequencies.

The contrast index (CI) is introduced to compare the enhancement effect of pear bruises at each spatial frequency. CI can quantitatively evaluate the image contrast, that is, the distinguishability of the bruised part relative to the whole part of the fruit. It needs to divide the pear to be detected into two parts, namely bruised tissue and sound tissue. Afterward, the ratio of the between-class variance to the total variance of the pixel intensity is calculated to obtain CI:

$$CI = \frac{N_x(\bar{x}-\bar{z})^2 + N_y(\bar{y}-\bar{z})^2}{\sum_{i=1}^{N_z}(z_i-\bar{z})^2} \quad (5)$$

where N_x, N_y, $N_z$are the number of pixels in the bruised, sound tissue and the whole region, respectively. And $\bar{x}$, $\bar{y}$ and $\bar{z}$ are the average strength of the bruised, sound tissue and the whole region, respectively. The value of CI is between 0 and 1, where a higher value indicates the better visibility and distinguishability of the bruised area.

Calculating the CI involves segmenting both bruised and sound areas, which can be challenging to achieve in AC images depicting mild bruises. On the contrary, RT images are more easily segmented due to contrast enhancement. Consequently, this study opts to

employ RT images rather than AC images to calculate CI for optimizing the spatial frequency. After removing the background by threshold segmentation, the Otsu threshold segmentation method (Otsu, 1979) was used to segment the bruise area to obtain the images of the bruised area, the sound area and the whole fruit area, and then the CI value was calculated according to Equation (5). The CI values under different spatial frequencies and different degrees of bruising were compared, and the optimal spatial frequency suitable for all degrees of bruising was selected for the next study.

2.5 Feature extraction and selection

Before using the machine learning algorithm to classify the images of pears, it is usually necessary to extract the features of the images, and use the extracted discriminant features to represent the images. Texture is one of the important features used to identify the object or region of interest in the image. Therefore, the texture features are also often applied to image classification in the fruit defect detection (Lu et al., 2021; Cai et al., 2022). Gray level co-occurrence matrix (GLCM) is a commonly used statistical method for image processing and texture analysis. It characterizes the texture of the image by calculating the frequency of pixel pairs with specific values and specific spatial relationships in the image to obtain GLCM, and then extracts statistical measures from the matrix. The Haralick features calculated based on GLCM are functions of distance and angle. In this study, 56 Haralick features with a distance of 1 were extracted in four directions (angles 0°, 45°, 90°135°) (Haralick et al., 1973). Therefore, 56 complete feature sets were extracted from each picture for bruise detection.

Feature selection is the process of selecting available feature subsets for prediction models. Feature selection serves to eliminate irrelevant or redundant features, resulting in a reduced feature set that can enhance model accuracy and decrease computation time. When dealing with limited data sets, feature selection can improve the generalization ability of machine learning models and mitigate overfitting occurrences. The Random Frog algorithm, originally introduced for gene selection, is a reversible jump Markov chain Monte Carlo (MCMC)-like algorithm Yun et al. (2013). This algorithm was used for feature selection. The process of feature selection includes feature subset search, feature subset evaluation and feature subset verification. Furthermore, choosing an appropriate stopping criterion can not only optimize the feature selection process but also reduce the overall selection time. The core idea of the random frog leaping method is to randomly select feature subsets. In this study, the performance evaluation and ranking of these subsets were conducted using the PLS-DA combined with cross-validation method. The outcomes of the feature selection were utilized to create a feature subset that will be employed for subsequent model classification.

2.6 Bruise classification algorithm

The pears with three degrees of bruising were classified. For each degree of bruising, the data set was randomly divided into training set and test set according to the ratio of 7:3.

This study developed three classification methods. LS-SVM is a variant of the standard Support Vector Machine (SVM). Unlike the latter, LS-SVM obtains the final decision function by solving linear equations instead of quadratic programming problems. As a result, it exhibits excellent generalization performance and requires lower computational cost (Suykens and Vandewalle, 1999). In this study, the radial basis function (RBF) kernel function was applied to the calculation of the LS-SVM, and the regularization parameters of the LS-SVM model were determined by ten-fold cross-validation. The purpose was to identify the parameter values that yield the best performance on the given dataset. PLS-DA is a supervised classification method, which was developed using the Partial Least Squares (PLS) algorithm initially designed for multivariate calibration (Wold et al., 2001). When employing the PLS-DA model for classification, it is crucial to ascertain the optimal number of latent variables for modeling. In this study, the number of latent variables in the PLS-DA model was determined based on the criterion of the smallest prediction error observed in the leave-one-out cross-validation. KNN is a widely employed machine learning algorithm for tackling supervised classification tasks. It functions by calculating the distance between various feature vectors and employs cross-validation to determine the most suitable value of K.

To address the variability introduced by data division, each of the aforementioned training instances is replicated 30 times. Each bruise degree and the overall samples were then modeled independently. The training was conducted using two distinct image inputs (AC and RT). Subsequently, a fair comparison was made between the outcomes obtained from the different image inputs and the three classifiers.

Three commonly used metrics are employed to assess the effectiveness of various classification models. These metrics include True Positive (TP), True Negative (TN), and Overall Accuracy (ACC). The TP and TN rates are computed as the ratios of accurately classified bruised and sound samples, respectively, to the total samples in their respective categories. ACC represents the proportion of all correctly classified samples to the total number of test samples. The aforementioned performance indicators are derived from the average values computed across thirty randomly partitioned datasets utilized for modeling.

The image preprocessing, feature extraction, and model training procedures were carried out using Matlab R2017a (The Mathworks, Inc., Natick, MA, USA).

3 Results and discussion

3.1 Performance of bruising detection based on different spatial frequencies

Figure 3 shows the basic image processing, including three-phase demodulation, background segmentation and frequency domain filtering. Using the three-phase demodulation method, the collected three SI images can be demodulated to obtain AC images and DC images. Background segmentation mainly used the

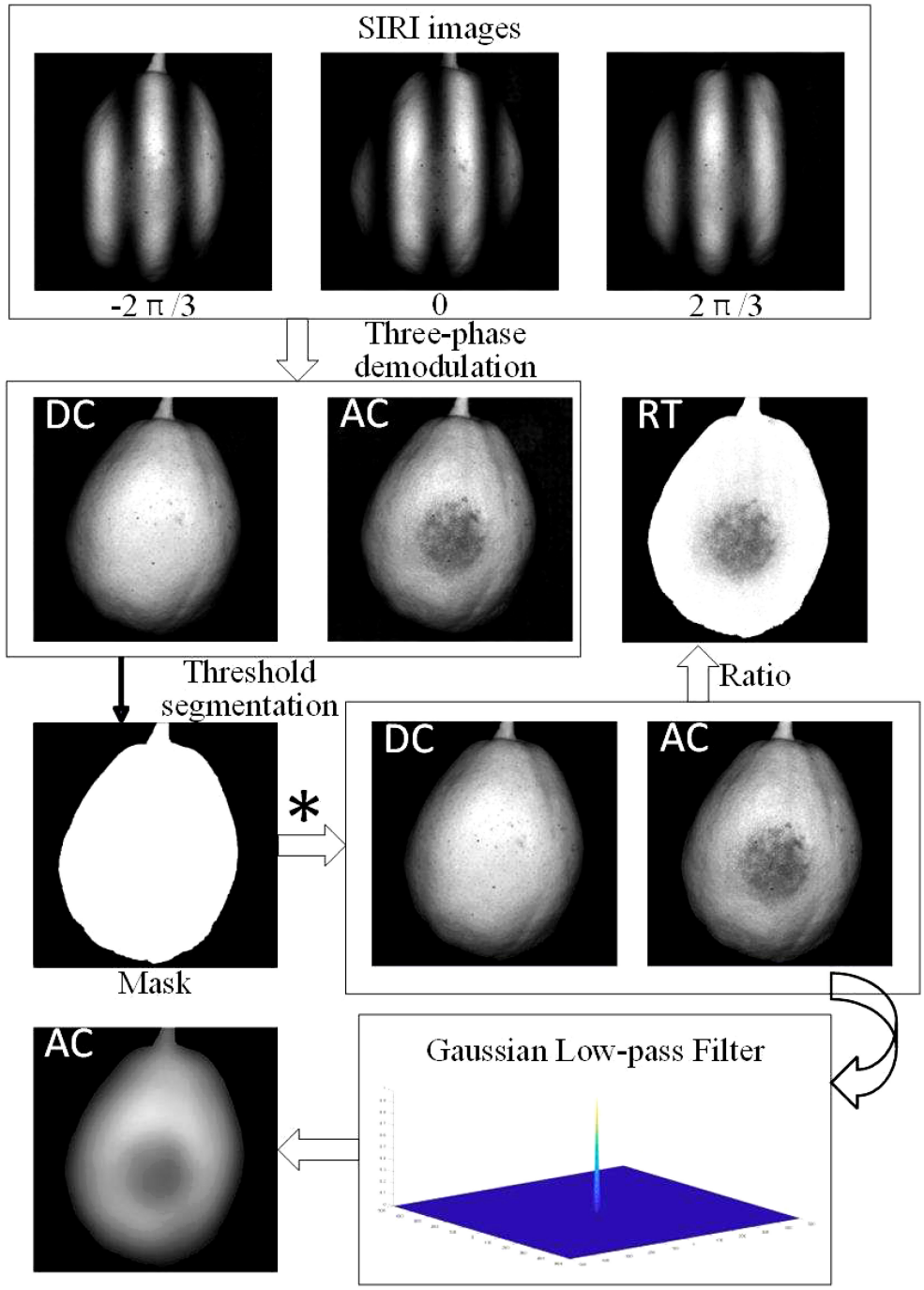


FIGURE 3
Three-Phase Demodulation and image processing. * Represents the dot product of Mask and DC or AC images.

threshold method to generate the pear area mask to remove the influence of the background on the bruise detection. Here, the DC image was used as a reference, and the mask was generated by manual threshold segmentation. The threshold is manually adjusted in a small increment to obtain the appropriate value, and the morphological operation is supplemented to generate the appropriate mask. Since the detection environment is stable, the value is used to generate a mask for all samples to segment the pear area from the image background. In addition, a Gaussian low-pass filter was used to denoise the AC images and enhance the bruise detection effect of the AC images. The processed AC images were used for the next step of image processing and classification.

Figure 4 displays the DC and RT images of three degrees of bruising (S1, S2, and S3) captured at different spatial frequencies. It should be noted that each spatial frequency produces a DC image, but all DC images remain basically the same. Upon visual inspection, it is observed that except for the spatial frequency of 50 cycles m^{-1}, RT images at different spatial frequencies can effectively identify the

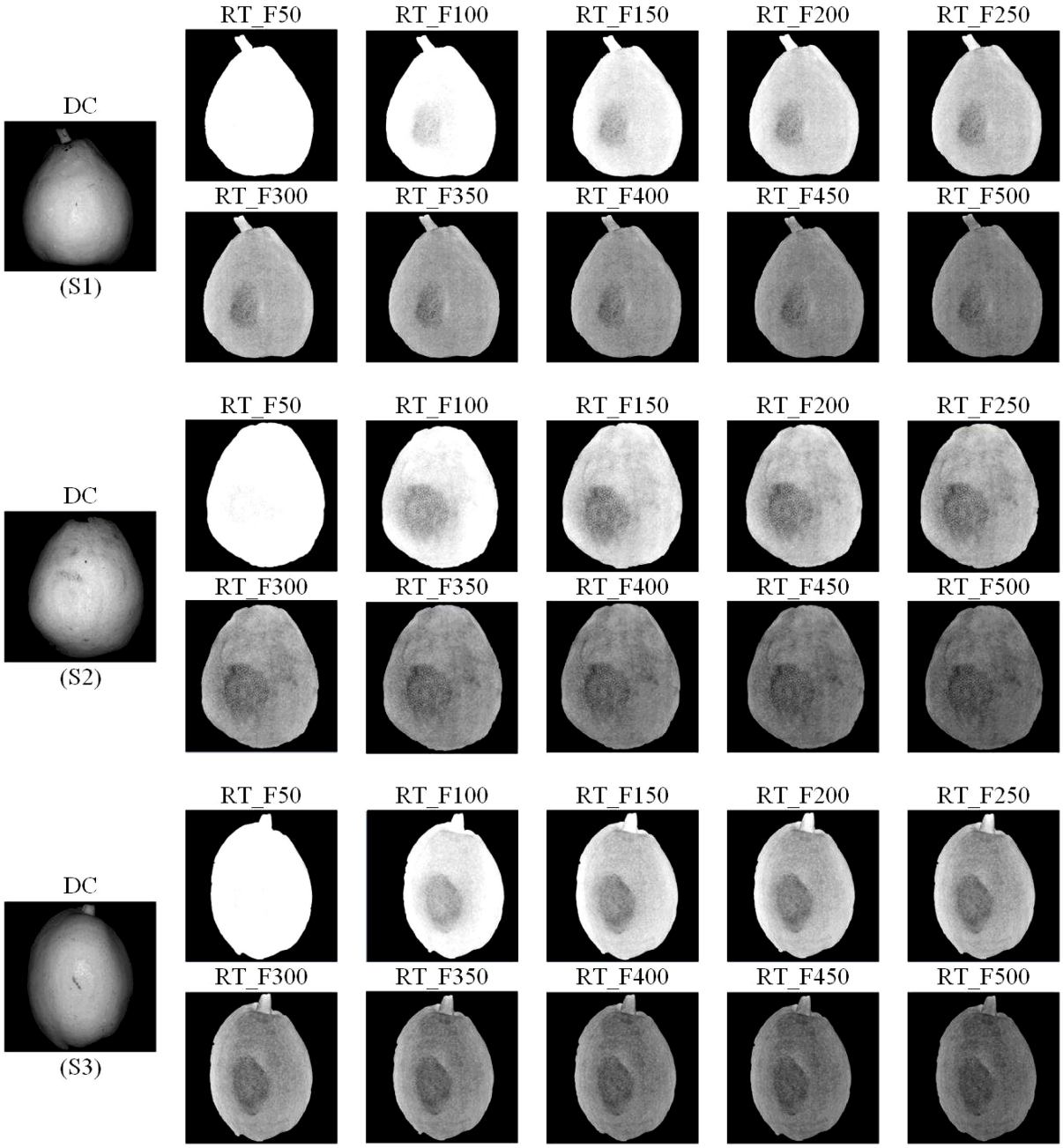


FIGURE 4
RT images and DC images for the mild (S1), moderate (S2) and severe bruised (S3) 'Korla' pears at the spatial frequencies of 50, 100, 150, 200, 250, 300, 350, 400, 450 and 500 cycles m^{-1}, respectively.

subcutaneous bruising area of pears, while the DC image, equivalent to the image under uniform illumination, does not show hidden bruises. In addition, the RT image led to a more uniform image background. Due to the curvature of the pear surface, the RT image has a positive effect on the correction of intensity distortions, which can greatly eliminate the influence of uneven illumination, while the DC image obviously shows a darker background edge. As the spatial frequency of the SI increases, the overall brightness of the RT image decreases. At higher spatial frequencies, as the overall brightness decreases, the bruise contrast decreases significantly. The darkening of RT images at high spatial frequencies can be attributed to the characteristics of SI. The SI attenuation rate at high spatial frequency is high, resulting in signal attenuation, so the AC image will be darkened. The brightness of the DC image at different frequencies does not change significantly, so the ratio image eventually darkens, as shown in Equation (4). In general, RT images at all frequencies except the lowest frequency achieved consistently good performance in detecting different fresh bruises on pears.

To further quantify the distinguishability of subcutaneous bruises in pears, the CI values were calculated for different bruise degrees at various spatial frequencies, as presented in Table 1. The table reveals a consistent pattern across different degrees of bruising: as the spatial frequency increases, the CI initially rises, reaching a peak at a certain frequency, and then gradually decreases. Indeed, except for the CI at spatial frequency of 50 cycles m^{-1}, bruises at low spatial frequencies are more distinguishable, which is in line with the visual observations.

Among them, S1 and S2 samples achieve the maximum CI at the spatial frequency of 150 cycles m^{-1}, whereas S3 samples reaches its highest CI at 100 cycles m^{-1}. Considering that mild bruises are more difficult to be detected, it is necessary to focus on the detectability of bruises in S1 and S2 samples. Moreover, it can be seen from the table that the CI values at 100 and 150 cycle m^{-1} in S3 are still at a high level. Hence, this study selected 150 cycle m^{-1} as the final spatial frequency for the subsequent bruise detection of all samples.

3.2 Image demodulation results

Figure 5 shows typical samples of three different degrees of bruising, all of which were detected immediately after the bruising occurred. These pears underwent varying degrees of bruising when exposed to different levels of static pressure. With naked eyes observation, bruises on pears are readily discernible in AC and RT images, whereas they are almost imperceptible in RGB and DC images. The bruised area appears as a darker region in the image. Nevertheless, it is impractical to determine the extent of bruising by relying solely on the grayscale values in this region. This limitation arises from the lack of discernible differences in intensity among the three distinct levels of bruising, particularly in RT image. The RT image clearly demonstrates effective image enhancement achieved by the ratio of AC to DC image. The contrast in the RT image is noticeably higher compared to the AC image, and it successfully eliminates artifacts resulting from the pear's surface color and irregular shape.

TABLE 1 Contrast indexes (CIs) obtained under for different spatial frequencies (cycles m^{-1}) three bruise degrees.

Bruise degree	50	100	150	200	250	300	350	400	450	500
S1	0.205	0.486	0.501	0.478	0.438	0.400	0.359	0.312	0.275	0.235
S2	0.382	0.611	0.625	0.536	0.486	0.433	0.384	0.343	0.309	0.276
S3	0.340	0.592	0.544	0.458	0.393	0.346	0.309	0.280	0.251	0.231

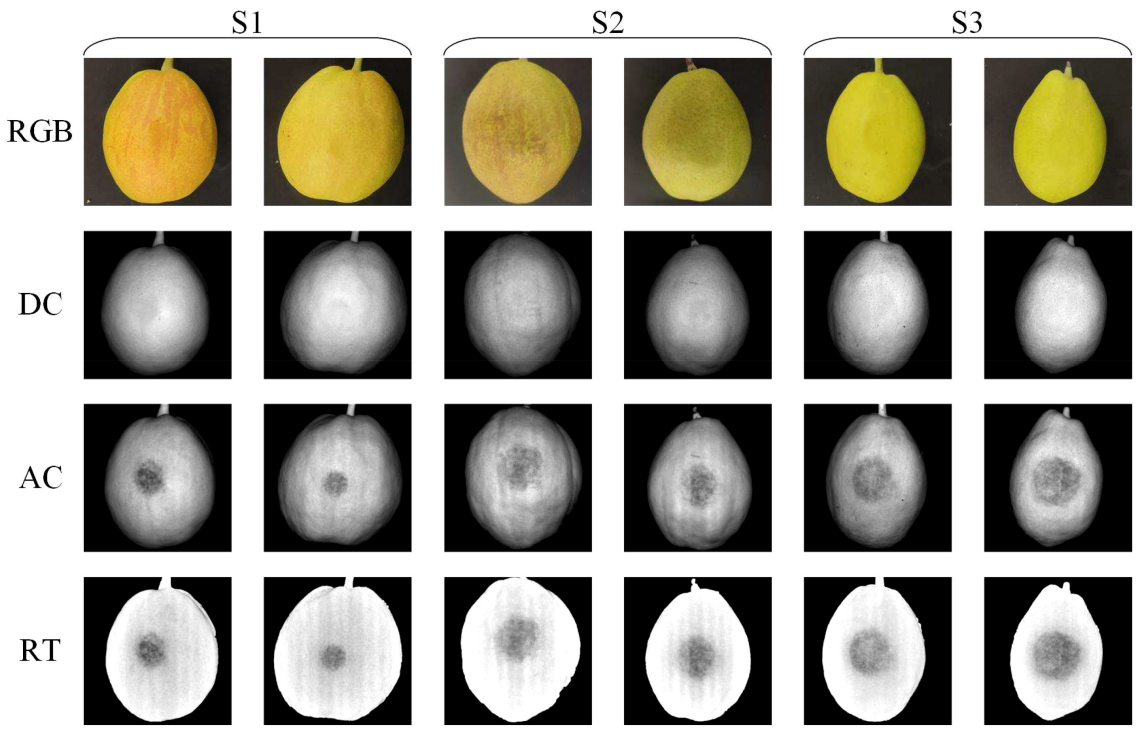


FIGURE 5
Typical RGB, DC, AC, RT images of mild (S1), moderate (S2) and severe (S3) bruise of 'Korla' pears.

It can be seen from the Figure 5 that pears at two maturity stages (green maturation period and yellow maturation period) can obtain good detection results. In addition, although some pears have red stripes, they have no effect on the final detection results including DC, AC and RT images. However, the irregular shape of pears does affect the detection results, mainly in DC and AC images, while RT images completely eliminate this negative influence.

3.3 Bruise classification

The classification outcomes of three classification models (LS-SVM, PLS-DA and KNN) when AC and RT images were employed as inputs for independent data were shown in Figures 6, 7, respectively. The diagrams illustrated that it was viable to employ visible LED SIRI technique to immediately detect the bruising on 'Korla' pears, resulting in a commendable level of detection accuracy. The detection accuracy of RT images under each classification model surpasses that of AC images, aligning with the observations made by visual inspection. The LS-SVM model exhibits both the highest detection accuracy and the greatest model stability. When compared to the PLS-DA and KNN models, LS-SVM demonstrates superior detection outcomes across three bruise severity levels and two image inputs. When AC images were used as input, the classification accuracy and stability of the LS-SVM model elevate as the degree of pear bruising. Notably, an overall classification accuracy exceeding 90% can still be achieved in the identification of mild bruising. From the perspective of ACC, the classification accuracy of LS-SVM, PLS-DA and KNN models increased with the increase of pear bruise degree. Among them, the accuracy of PLS-DA in detecting samples with severe bruise degree was close to that of LS-SVM model, but its stability was still not as good as the latter. The KNN model also achieved 92.3% ACC, but its stability is far less than the former two. In actual production, the degree of bruising of pears is not the same, which is related to the environment of pears in production and transportation. Therefore, the overall detection accuracy of different degrees of bruising may be more in line with the actual situation. Although the LS-SVM model achieves high accuracy and stability in the detection of samples with a single degree of bruising, it's not very outstanding in the detection of bruises in mixed samples with three degrees of bruising due to only 85.6% of sound fruit recognition accuracy. Therefore, AC images may not be suitable for bruise detection of pears in commercial production.

When RT images were used as input, the three classification models show excellent performance in bruise detection accuracy, and were superior to AC images in terms of detection accuracy and model stability. Moreover, according to the error bar, the stability of the LS-SVM model is still higher than that of the other two models. For individual and combined samples with different degrees of

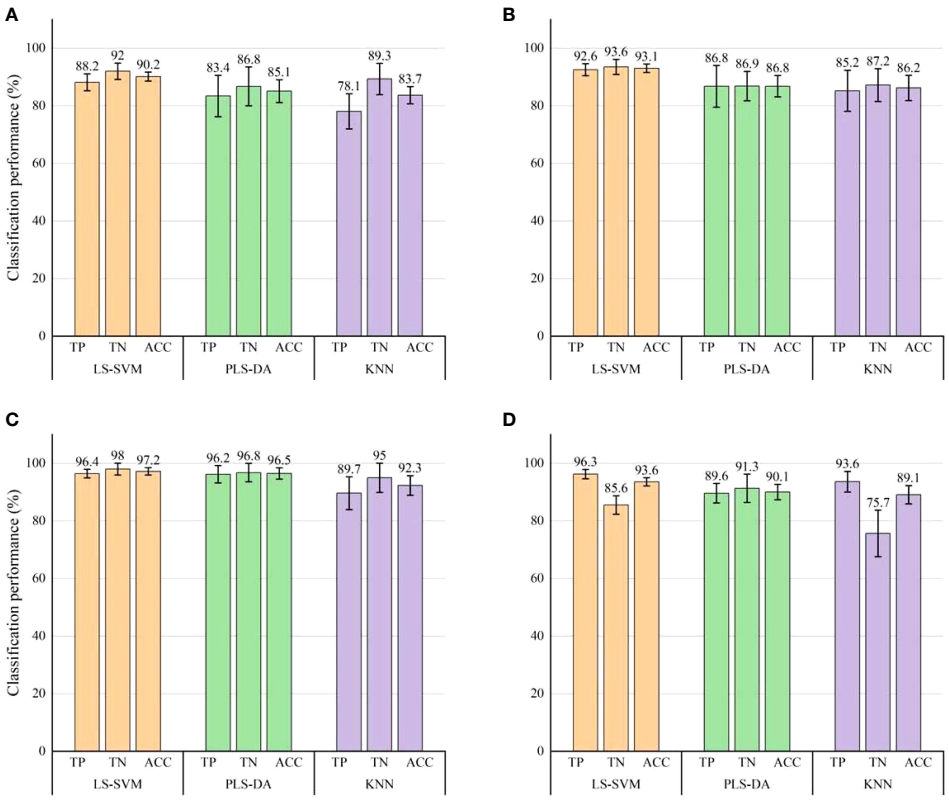


FIGURE 6
Classification results for bruise detection by using three classification models with AC images. **(A)** Classification results of mild bruises (S1). **(B)** Classification results of moderate bruises (S2). **(C)** Classification results of severe bruises (S3). **(D)** Overall classification results of the three levels of bruising. Error bars on the chart signifies the corresponding standard errors of the evaluation index derived from 30 modeling instances.

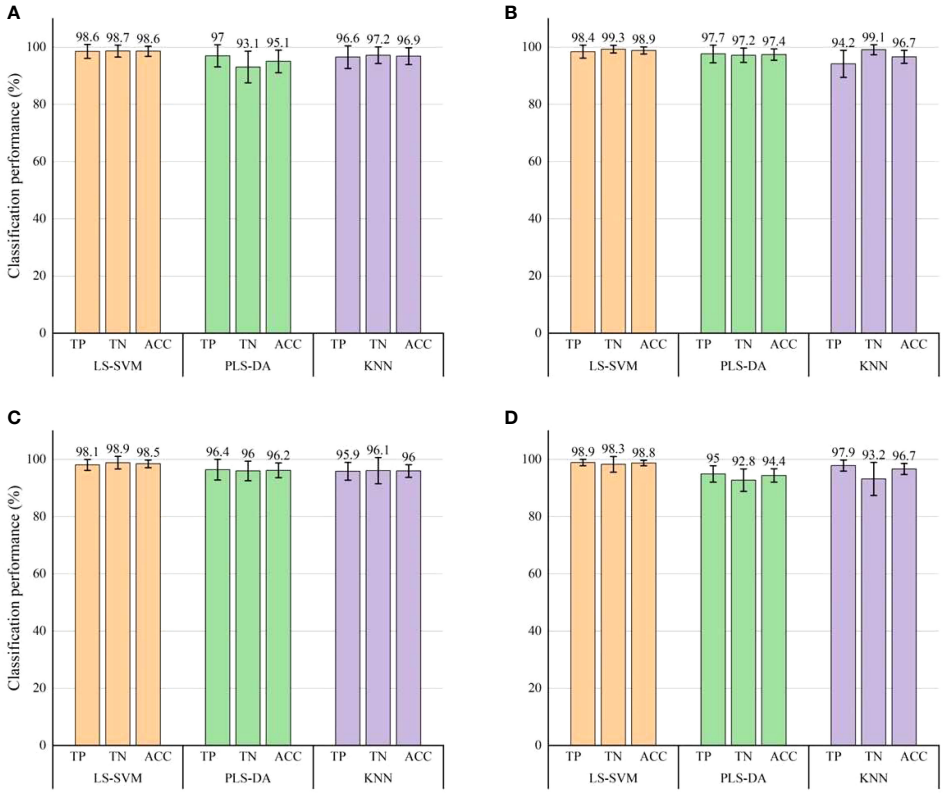


FIGURE 7
Classification results for bruise detection by using three classification models with RT images. **(A)** Classification results of mild bruises (S1). **(B)** Classification results of moderate bruises (S2). **(C)** Classification results of severe bruises (S3). **(D)** Overall classification results of the three levels of bruising. Error bars on the chart signifies the corresponding standard errors of the evaluation index derived from 30 modeling instances.

bruising, the three evaluation indexes (TP, TN and ACC) of the LS-SVM model all exceeded 98%. Notably, when the classification model was employed for identifying bruises in RT images, it consistently maintains a high level of accuracy in detecting bruises of varying severities, with little fluctuation. This shows that compared with AC images, the detection effect of RT images is less correlated with the degree of pear bruising. The detection accuracy of RT images in each degree of bruise was greater than the best result of AC images in detecting bruises (severe bruises). Therefore, it is a better choice to use RT images as the basis for pear bruising detection. Especially, the detection performance is still very good when the RT image performs mixed detection of pears with different bruising degrees. Hence, RT image was proved to be a more favorable option for detecting bruises on pears with varying degrees. It is feasible to use RT images for bruising detection of 'Korla' pears in practice.

In feature selection, different feature subsets will be generated according to the different division of sample sets. In this study, based on the random frog feature selection algorithm, the optimal ten features were selected as feature subsets. With the division of each data set, the number of times each feature is selected as a feature subset is counted, as shown in Figure 8. Ten features with the most selected times are selected to establish a new feature subset. It can be seen from the figure that the feature subsets of different degrees of bruising are not the same, and the feature subsets of AC images and RT images are also quite different. Among them, the feature subsets of AC images with different degrees of bruising are quite different, and there are few common features, while there are many common features for RT images, which further proves that the stability of RT image for detection of bruising is higher than that of AC image. In addition, for the mixed data sets of three bruising degrees, the feature subset of the AC image only contains the most frequently selected features in the independent data sets of different bruising degrees, while the RT image contains many common features, which indicates that it is easier to obtain the best subset of features from the feature set of the RT image to achieve the highest classification accuracy. From the perspective of detection accuracy, the accuracy of the classification model with AC image as input will increase with the increase of the degree of bruising, while the classification model with RT image as input has little difference in accuracy, which is consistent with the results of feature selection.

Table 2 summarizes the classification accuracy of bruised pears by three kinds of models established based on ten features. These features were selected based on the above highest frequency. It can be noted that different models have varying classification accuracy for inputs with different degrees of bruising and AC/RT images. The classification accuracy of the model with RT image as input is still higher than that of AC image regardless of type of model, and the

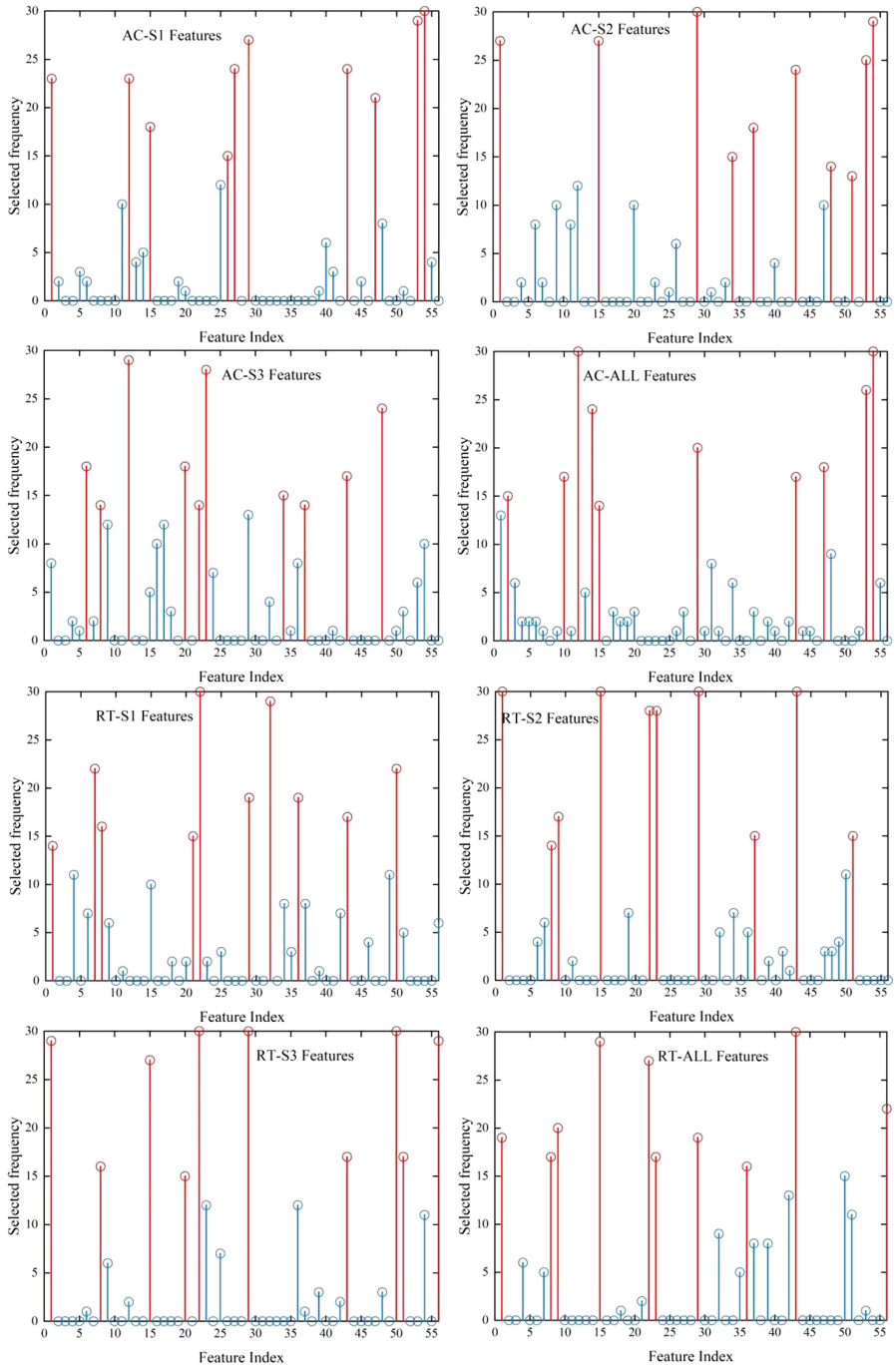


FIGURE 8
The feature selection results of independent data set and mixed data set of AC image and RT image with mild (S1), moderate (S2) and severe (S3) bruises. The first ten most discriminative features are selected by the random frog feature selection algorithm, and the frequency of each feature selected when the data set is randomly divided for thirty times is counted. The top ten features according to selected frequency are highlighted in red.

LS-SVM model is the optimal classification model. The overall detection accuracies of slight, moderate, severe and mixed degree of bruising were 99%, 98.11%, 98.44% and 98.64%, respectively. By using the feature subset with the highest frequency, the LS-SVM model improved the effect of detecting mild bruises when RT image was used as input. Further examining the ten selected texture features, it is found that they mainly come from the Angular Second Momen, sum entropy, entropy, and maximum correlation coefficient in different directions, indicating they, combined with RT image, are important for detecting 'Korla' pear bruises.

4 Conclusion

This study successfully demonstrated the feasibility of low-cost visible-LED SIRI technique for the early detection of varying

TABLE 2 The classification accuracy (%) of bruised pears by three kinds of models established based on ten features with the highest frequency of selection.

Input	Degree	LS-SVM			PLS-DA			KNN		
		TP	TN	ACC	TP	TN	ACC	TP	TN	ACC
AC	S1	89.33	92.33	90.83	84.78	88.78	86.78	77.33	89.89	83.61
	S2	92.22	94.00	93.11	85.78	89.22	87.50	85.78	90.89	88.33
	S3	96.11	98.56	97.33	94.56	90.00	92.28	89.67	96.44	93.06
	ALL	96.59	88.00	94.44	90.04	94.33	91.11	93.41	79.67	89.97
RT	S1	99.00	99.00	99.00	93.44	92.56	93.00	95.67	97.44	96.56
	S2	97.33	98.89	98.11	88.44	93.56	91.00	98.44	98.44	98.44
	S3	98.00	98.89	98.44	99.44	98.89	99.17	96.44	95.78	96.11
	ALL	99.11	97.22	98.64	92.26	92.89	92.42	96.78	85.44	93.94

ALL refers to a collection of mild (S1), moderate (S2) and severe (S3) bruising samples.

degrees of subcutaneous bruising in 'Korla' pears. The 150 cycles m^{-1} was determined as the optimal structural illumination spatial frequency. For detection of three degrees of bruised pears, RT and AC images were significantly superior to DC images, and RT image was best due to the ability of enhanced image contrast and brightness unevenness correction. Texture features can serve as important features for classifying bruised and sound pears and random frog was an effective texture feature optimization algorithm. Among three types of texture feature models (LS-SVM, PLS-DA and KNN models), the LS-SVM model exhibited superior detection performance with the highest detection accuracy and stability, regardless of single bruising degree classification or mixed bruising degree classification. The LS-SVM model established using ten appropriate features extracted from RT images achieved classification accuracies of 98.6%, 98.9%, 98.5%, and 98.8% for mild, moderate, severe and mixed bruises, respectively, indicating the outstanding ability of the proposed methodology in detecting the bruising of pear fruit in this study. Subsequent study should improve the hardware system and algorithms so that this low-cost SIRI technique can be implemented for online detection of pear bruising. Furthermore, the capacity for the early bruising detection of other thin-skinned fruit (e.g. apple and peach) should be also assessed to expand the application of this technology.

Data availability statement

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

Author contributions

MM: Data curation, Methodology, Writing – original draft. ZC: Writing – review & editing, Validation. XZ: Writing – review & editing. CS: Writing – review & editing. JZ: Writing – review & editing. HP: Funding acquisition, Writing – review & editing. JL: Funding acquisition, Writing – review & editing, Supervision. RS: Writing – review & editing, Validation. WZ: Writing – review & editing, Validation.

Funding

The author(s) declare financial support was received for the research, authorship, and/or publication of this article. Authors are grateful for the National Natural Science Foundation of China (31972152), Jiangsu Province and Education Ministry Co-sponsored Synergistic Innovation Center of Modern Agricultural Equipment (XTCX2001) and the Bingtuan Science and Technology Program (2022DB004) for providing funding for the study.

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 author(s) declared that they were an editorial board member of Frontiers, at the time of submission. This had no impact on the peer review process and the final decision.

Publisher's note

References

Azadbakht, M., Vahedi Torshizi, M., and Mahmoodi, M. J. (2019a). The use of CT scan imaging technique to determine pear bruise level due to external loads. *Food Sci. Nutr.* 7 (1), 273–280. doi: 10.1002/fsn3.882

Azadbakht, M., Vahedi Torshizi, M., and Mahmoodi, M. J. (2019b). The relation of pear volume and it's bruised volume by CT scan imaging. *J. Food Measurement Characterization* 13 (2), 1089–1099. doi: 10.1007/s11694-018-00024-0

Bian, H., Shi, P., and Tu, P. (2020). Determination of physicochemical quality of bruised apple using dielectric properties. *J. Food Measurement Characterization* 14 (5), 2590–2599. doi: 10.1007/s11694-020-00505-1

Cai, Z., Huang, W., Wang, Q., and Li, J. (2022). Detection of early decayed oranges by structured-illumination reflectance imaging coupling with texture feature classification models. *Front. Plant Sci.* 13. doi: 10.3389/fpls.2022.952942

Celik, H. K. (2017). Determination of bruise susceptibility of pears (Ankara variety) to impact load by means of FEM-based explicit dynamics simulation. *Postharvest Biol. Technol.* 128, 83–97. doi: 10.1016/j.postharvbio.2017.01.015

Fu, X., and Wang, M. (2022). Detection of early bruises on pears using fluorescence hyperspectral imaging technique. *Food Analytical Methods* 15 (1), 115–123. doi: 10.1007/s12161-021-02092-3

Haralick, R. M., Shanmugam, K., and Dinstein, I. H. (1973). Textural features for image classification. *IEEE Trans. systems man cybernetics* 6), 610–621. doi: 10.1109/TSMC.1973.4309314

He, X., Jiang, X., Fu, X., Gao, Y., and Rao, X. (2018). Least squares support vector machine regression combined with Monte Carlo simulation based on the spatial frequency domain imaging for the detection of optical properties of pear. *Postharvest Biol. Technol.* 145, 1–9. doi: 10.1016/j.postharvbio.2018.05.018

Hussein, Z., Fawole, O. A., and Opara, U. L. (2019). Bruise damage susceptibility of pomegranates (Punica granatum, L.) and impact on fruit physiological response during short term storage. *Scientia Hortic.* 246, 664–674. doi: 10.1016/j.scienta.2018.11.026

Khurnpoon, L., and Siriphanich, J. (2011). Change in fruit quality and cell wall polysaccharides in bruised papaya cultivars 'Khak dum' and 'Holland'. *Int. Conf. Postharvest Unlimited* 945, 381–389. doi: 10.17660/ActaHortic.2012.945.52

Kim, G., Kim, G. H., Park, J., Kim, D. Y., and Cho, B. K. (2014). Application of infrared lock-in thermography for the quantitative evaluation of bruises on pears. *Infrared Phys. Technol.* 63, 133–139. doi: 10.1016/j.infrared.2013.12.015

Lee, W. H., Kim, M. S., Lee, H., Delwiche, S. R., Bae, H., Kim, D. Y., et al. (2014). Hyperspectral near-infrared imaging for the detection of physical damages of pear. *J. Food Eng.* 130, 1–7. doi: 10.1016/j.jfoodeng.2013.12.032

Li, J., Lu, Y., and Lu, R. (2023). Detection of early decay in navel oranges by structured-illumination reflectance imaging combined with image enhancement and segmentation. *Postharvest Biol. Technol.* 196, 112162. doi: 10.1016/j.postharvbio.2022.112162

Li, J., Lu, Y., and Lu, R. (2024). Identification of early decayed oranges using structured-illumination reflectance imaging coupled with fast demodulation and improved image processing algorithms. *Postharvest Biol. Technol.* 207, 112627. doi: 10.1016/j.postharvbio.2023.112627

Lu, Y., Li, R., and Lu, R. (2016). Structured-illumination reflectance imaging (SIRI) for enhanced detection of fresh bruises in apples. *Postharvest Biol. Technol.* 117, 89–93. doi: 10.1016/j.postharvbio.2016.02.005

Lu, Y., and Lu, R. (2017). Non-destructive defect detection of apples by spectroscopic and imaging technologies: a review. *Trans. ASABE* 60 (5), 1765–1790. doi: 10.13031/trans.12431

Lu, Y., and Lu, R. (2019). Structured-illumination reflectance imaging for the detection of defects in fruit: Analysis of resolution, contrast and depth-resolving features. *Biosyst. Eng.* 180, 1–15. doi: 10.1016/j.biosystemseng.2019.01.014

Lu, Y., Lu, R., and Zhang, Z. (2021). Detection of subsurface bruising in fresh pickling cucumbers using structured-illumination reflectance imaging. *Postharvest Biol. Technol.* 180, 111624. doi: 10.1016/j.postharvbio.2021.111624

Mei, M., and Li, J. (2023). An overview on optical non-destructive detection of bruises in fruit: technology, method, application, challenge and trend. *Comput. Electron. Agric.* 213, 108195. doi: 10.1016/j.compag.2023.108195

Opara, U. L., and Pathare, P. B. (2014). Bruise damage measurement and analysis of fresh horticultural produce-A review. *Postharvest Biol. Technol.* 91, 9–24. doi: 10.1016/j.postharvbio.2013.12.009

Otsu, N. (1979). A threshold selection method from gray-level histograms. *IEEE Trans. systems man cybernetics* 9 (1), 62–66. doi: 10.1109/TSMC.1979.4310076

Polat, R., Aktas, T., and Ikinci, A. (2012). Selected mechanical properties and bruise susceptibility of nectarine fruit. *Int. J. Food Properties* 15 (6), 1369–1380. doi: 10.1080/10942912.2010.498546

Razavi, M. S., Asghari, A., Azadbakh, M., and Shamsabadi, H. A. (2018). Analyzing the pear bruised volume after static loading by Magnetic Resonance Imaging (MRI). *Scientia Hortic.* 229, 33–39. doi: 10.1016/j.scienta.2017.10.011

Razavi, M. S., Golmohammadi, A., Sedghi, R., and Asghari, A. (2020). Prediction of bruise volume propagation of pear during the storage using soft computing methods. *Food Sci. Nutr.* 8 (2), 884–893. doi: 10.1002/fsn3.1365

Schreiber, H., and Bruning, J. H. (2007). "Phase shifting interferometry," in *Optical Shop Testing, 3rd Edition*. Ed. D. Malacara (Hoboken, New Jersey: John Wiley and Sons), 547–666. doi: 10.1002/9780470135976.ch14

Sun, Y., Lu, R., Lu, Y., Tu, K., and Pan, L. (2019). Detection of early decay in peaches by structured-illumination reflectance imaging. *Postharvest Biol. Technol.* 151, 68–78. doi: 10.1016/j.postharvbio.2019.01.011

Suykens, J. A., and Vandewalle, J. (1999). Least squares support vector machine classifiers. *Neural Process. Lett.* 9, 293–300. doi: 10.1023/A:1018628609742

Tian, X., Liu, X., He, X., Zhang, C., Li, J., and Huang, W. (2023). Detection of early bruises on apples using hyperspectral reflectance imaging coupled with optimal wavelengths selection and improved watershed segmentation algorithm. *Journal of the Science of Food and Agriculture* 103 (13), 6689–6705. doi: 10.1002/jsfa.12764

Tian, X., Zhang, C., Li, J., Fan, S., Yang, Y., and Huang, W. (2021). Detection of early decay on citrus using LW-NIR hyperspectral reflectance imaging coupled with two-band ratio and improved watershed segmentation algorithm. *Food Chemistry* 360, 130077. doi: 10.1016/j.foodchem.2021.130077

Wold, S., Sjöström, M., and Eriksson, L. (2001). PLS-regression: a basic tool of chemometrics. *Chemometrics intelligent Lab. Syst.* 58 (2), 109–130. doi: 10.1016/S0169-7439(01)00155-1

Wu, J., Li, F., Ge, Y., and Hu, R. (2013). Measurement of contact pressure of Korla pear under compression and bruising predication using finite element analysis. *Trans. Chin. Soc. Agric. Eng.* 29 (6), 261–266.

Yun, Y. H., Li, H. D., Wood, L. R., Fan, W., Wang, J. J., Cao, D. S., et al. (2013). An efficient method of wavelength interval selection based on random frog for multivariate spectral calibration. *Spectrochimica Acta Part A: Mol. Biomolecular Spectrosc.* 111, 31–36. doi: 10.1016/j.saa.2013.03.083

Zeng, X., Miao, Y., Ubaid, S., Gao, X., and Zhuang, S. (2020). Detection and classification of bruises of pears based on thermal images. *Postharvest Biol. Technol.* 161, 111090. doi: 10.1016/j.postharvbio.2019.111090

Zhou, Y., Mao, J., Wu, D., Liu, T., Zhao, Y., Zhou, W., Chen, Z., et al. (2019). Nondestructive early detection of bruising in pear fruit using optical coherence tomography. *Horticultural Science and Technology* 37 (1), 140–150. doi: 10.12972/kjhst.20190013

frontiers | Frontiers in Plant Science

TYPE Original Research
PUBLISHED 18 December 2023
DOI 10.3389/fpls.2023.1260644

OPEN ACCESS

EDITED BY
Te Ma,
Nagoya University, Japan

REVIEWED BY
Tanima Bhattacharya,
Lincoln University College, Malaysia
Panmanas Sirisomboon,
King Mongkut's Institute of Technology Ladkrabang, Thailand

*CORRESPONDENCE
Shimeles Tilahun
shimeles@kangwon.ac.kr
Cheon Soon Jeong
jeongcs@kangwon.ac.kr

†These authors have contributed equally to this work

RECEIVED 18 July 2023
ACCEPTED 01 December 2023
PUBLISHED 18 December 2023

CITATION
Baek MW, Choi HR, Hwang IG, Tilahun S and Jeong CS (2023) Prediction of tannin content and quality parameters in astringent persimmons from visible and near-infrared spectroscopy.
Front. Plant Sci. 14:1260644.
doi: 10.3389/fpls.2023.1260644

Prediction of tannin content and quality parameters in astringent persimmons from visible and near-infrared spectroscopy

Min Woo Baek[1,2†], Han Ryul Choi[3†], In Geun Hwang[4], Shimeles Tilahun[5,6*] and Cheon Soon Jeong[1,2*]

[1]Department of Horticulture, Kangwon National University, Chuncheon, Republic of Korea, [2]Interdisciplinary Program in Smart Agriculture, Kangwon National University, Chuncheon, Republic of Korea, [3]National Institute of Horticultural and Herbal Science, Rural Development Administration, Wanju-gun, Republic of Korea, [4]Frusen Co., LTD, Yongin, Republic of Korea, [5]Agriculture and Life Science Research Institute, Kangwon National University, Chuncheon, Republic of Korea, [6]Department of Horticulture and Plant Sciences, Jimma University, Jimma, Ethiopia

Introduction: Tannin content and postharvest quality characteristics of persimmon fruit are often determined by the destructive analysis that consumes time, does not allow the acquisition of data from the same fruit continuously, and requires expensive high-performance equipment. This research was done to investigate the potential for non-destructive estimation of astringency and quality parameters in persimmon fruit based on visible/near-infrared (VNIR) spectra.

Methods: VNIR spectra readings, the reference tannin content, and quality parameters were measured from fruits of "Cheongdo-Bansi" and "Daebong" persimmon cultivars at harvest and throughout the ripening/deastringency period. The spectra readings from half of the total fruit were utilized for the calibration set, while the other half readings were used for the prediction set. To develop models correlating the spectra data to the measured reference parameters data, the partial least square regression (PLSR) method was utilized.

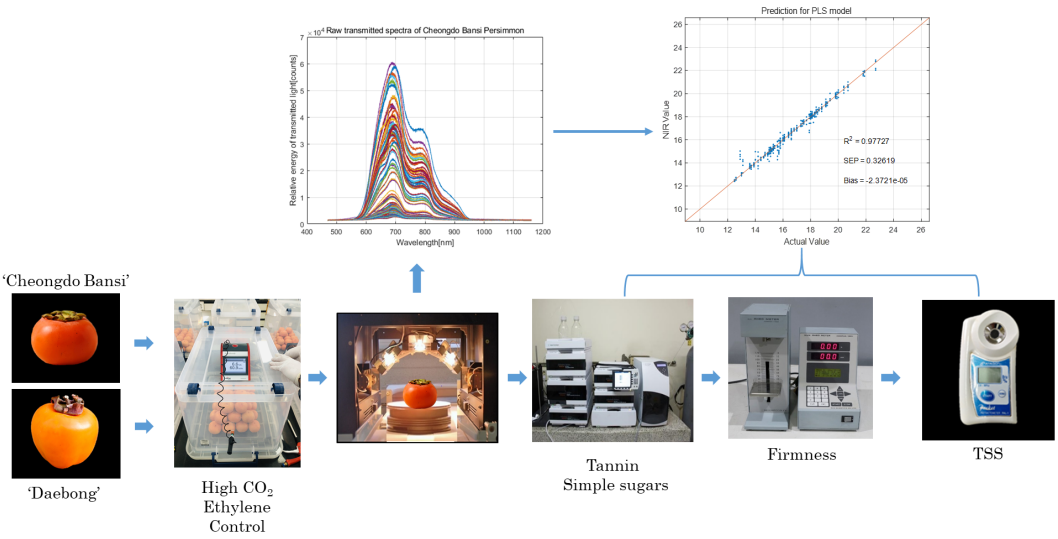
Results and discussion: In the case of 'Daebong', the coefficients of determination (R^2) between VNIR spectra and the actual measured values of TSS, firmness, simple sugars, and tannin content were (0.95, 0.94, 0.96, and 0.96) and (0.93, 0.89, 0.96, and 0.93), for the calibration and prediction sets, respectively. Similarly, the R^2-values of (0.86, 0.93, 0.79, and 0.81) and (0.83, 0.91, 0.75, and 0.75) were recorded in 'Cheongdo-Bansi' for the calibration and prediction sets, respectively. Additionally, the acquired data were divided into two sets in a 3:1 ratio to develop predictive models and to validate the models in multiple regressions. PLSR models were developed in multiple regression to estimate the tannin content of both cultivars from firmness and simple sugars with R^2-values of 0.83 and 0.79 in 'Cheongdo-Bansi' for the calibration and prediction sets, respectively, whereas, R^2-values of 0.80 and 0.84 were recorded in 'Daebong' for the calibration and

10.3389/fpls.2023.1260644

prediction sets, respectively. The overall findings of this study showed the possibility of using VNIR spectra for the prediction of postharvest quality and tannin contents from intact persimmon fruit with quick, chemical-free, and low-cost assessment methods. Also, the multiple regression using physicochemical parameters could fairly predict the tannin content in persimmon fruit though destructively but save time and low-cost.

KEYWORDS

persimmon, astringency, tannin, non-destructive PLS model, Vis/NIR spectra



GRAPHICAL ABSTRACT

1 Introduction

Persimmon (*Diospyros kaki* Thunb.) probably originated in China and has been mainly grown in China, Korea, and Japan as a relevant food source from prehistoric times (Parfitt et al., 2015). In 2021, China, the Republic of Korea, and Japan contributed 96% of the world's persimmon production (FAOSTAT, 2021). According to FAOSTAT (2021), persimmon production in the Republic of Korea was 200,610 tons from the total Asia and world production of 4.16 and 4.33 million tons, respectively. Persimmon fruit classifies as either astringent or non-astringent, and it is a delicious and healthy fruit rich in vitamins, minerals, and antioxidants which are associated with numerous health benefits (Park et al., 2017; Das and Eun, 2021). 'Cheongdo-Bansi' and 'Daebong' are commercially important astringent persimmon cultivars that are commonly grown in the Republic of Korea due to their adaptability to the environment and excellent taste and texture (Park et al., 2017; Park et al., 2019). Tannins are polyphenol compounds with a high molecular weight that cause astringency due to their large hydroxyl phenolic groups (Cortés et al., 2017). The soluble tannins gradually turn into insoluble tannins as the fruit ripens and the fruit become less astringent (Noypitak et al., 2015; Cortés et al., 2017). In non-astringent persimmon, soluble tannin is reduced naturally during ripening, while in astringent persimmon, a high level of soluble tannin is maintained when it is not fully ripe and fruits cannot be eaten during the commercial harvest stage because of their higher levels of soluble tannins (Yamada et al., 2002; Akagi et al., 2009; Das and Eun, 2021). Fruits of both 'Cheongdo-Bansi' and 'Daebong' cultivars, however, undergo rapid softening after harvest, and by the time astringency is low enough to be palatable, the fruits become too soft. Conversely, firm textured fruits which are suitable for distribution are astringent. This astringency can cause a dry or puckering sensation in the mouth that can be unpleasant (Das and Eun, 2021). Hence, it requires rapid ripening or removing the astringency from persimmons for agreeable palatability.

However, many studies were reported to achieve fast removal of astringency from persimmon including ethylene (Park et al., 2017; Park et al., 2019), high concentrations of CO_2 (Arnal and Del Río, 2003; Salvador et al., 2007), ethanol (Ortiz et al., 2005), high (Ben-Arie and Sonego, 1993) and freezing (Das and Eun, 2021) temperatures treatments. Treatments with ethylene and high concentration of CO_2 (high CO_2) are the most widely used commercial techniques that promote fast ripening and astringency removal, respectively (Cortés et al., 2017; Park et al., 2017). Ethylene treatment causes rapid expression of ripening-related genes (Park et al., 2017; Park et al., 2019), and exposing the fruit to a high CO_2 promotes the accumulation of acetaldehyde due to anaerobic respiration in the fruit. The soluble tannins become insoluble as they react with the acetaldehyde and the astringency is thus eliminated (Cortés et al., 2017).

Firmness is the main difference between the persimmon fruit deastringed by the treatment with ethylene and the high CO_2. Fruit treated with ethylene becomes softer, and acquires a jelly-like consistency which is difficult to distribute. Yet, some consumers prefer the taste and store it in a deep freezer for future use after fully ripens. On the other hand, the firm texture of the fruit is maintained during deastringency with the high CO_2, appreciated by the industry and consumers in its suitability for distribution (Munera et al., 2017a). Therefore, optimum ripening and astringency removal are required to avoid loss of fruit quality caused by high concentration or long treatment and residual astringency due to low concentration or short treatment of ethylene or high CO_2 (Arnal et al., 2008; Novillo et al., 2014; Park et al., 2017). Hence, it is important to measure tannin contents during the treatment periods to ensure optimum ripening and deastringency.

The common methods used to measure the changes in tannin contents during ripening and astringency removal are usually destructive and thus the same fruit cannot be monitored continuously. The analysis also requires expensive high-performance equipment and consumes time. Therefore, having a reliable, low-cost, fast, and easy-to-implement method for tannin determination in persimmons is a useful tool for astringency management during postharvest handling and distribution. Predictive models developed by using visible and near-infrared (VNIR) spectroscopy and color variables are among the most common techniques currently used for the prediction of secondary metabolites such as lycopene and β-carotene in tomatoes and α-solanine and α-chaconine in potatoes (Tilahun et al., 2018; Tilahun et al., 2020). The interaction between VNIR range spectra and the organic molecules that make up a compound helps to obtain qualitative and quantitative information from the spectra (Pasquini, 2003; Tilahun et al., 2020).

Cortés et al. (2016) predicted the internal quality (combination of total soluble solids (TSS), firmness, and flesh color) of mango with VNIR reflection spectroscopy. Noypitak et al. (2015) also developed PLSR models to evaluate tannin content in astringent 'Xichu' persimmon and recommended NIR interactance spectroscopy for optimal prediction of soluble tannin content. In addition, Cortés et al. (2017) also reported the possibility of determining astringency through reflectance VNIR spectra at selected points in intact and half-cut 'Rojo Brillante' persimmon fruit. Similarly, Son et al. (2009) predicted sugar contents in a sweet persimmon using reflectance spectra. Most of the previous works if not all, used CO_2 treatment to remove the astringency of persimmon during nondestructive estimations. However, more samples at different levels of astringency and softening are needed to ascertain the prediction power of astringency and ripening quality to fulfill the demands of both consumers and the industry. Thus, this study included the treatment with ethylene or high CO_2 and untreated control of the intact 'Daebong' and 'Cheongdo-Bansi' persimmon fruits. This work determines the possibility of nondestructive estimation of astringency and quality parameters including TSS, firmness, and simple sugars by using VNIR spectroscopy in transmittance mode, in combination with a multivariate analysis technique, to predict the changes in quality and tannin content of persimmon fruits during ripening and deastringency.

2 Materials and methods

2.1 Plant material, treatments and storage at ambient condition

Astringent persimmon fruits (*Diospyros kaki* Thunb. 'Cheongdo-Bansi' and 'Daebong') were harvested from Jeollanamdo, Yeoungham, Korea on 28 Sept. 2022. After harvest, 150 uniform fruits free of external damage were selected from each cultivar. Within 12 hours of harvest, the fruits were then brought to the postharvest laboratory at the Department of Horticultural Sciences, Kangwon National University, Korea. After keeping at ambient condition for 3 hours to remove field heat, uniform fruits free of defects were carefully reselected and divided into three groups (control, ethylene treatment, and high CO_2 treatment, 40 fruits each) for both cultivars. So, 120 fruits of each cultivar were used for the experiment. The treatment groups were treated separately with 100 mg kg^{-1} ethylene (Park et al., 2017) and 95% CO_2 (Arnal et al., 2008) for 24 h in a sealed 62 L container at 22°C. The control fruits were treated under similar conditions without ethylene and CO_2 treatment. The fruits were characterized as 117.8 ± 1.96 and 271.5 ± 1.73 g of fresh weight, 20.29 ± 0.8 and 19.26 ± 0.9 N of firmness, 16.93 ± 0.5 and 17.48 ± 0.4% of TSS, and 4.68 ± 0.2 and 4.98 ± 0.3 g kg^{-1} of soluble tannin at harvest for 'Cheongdo-Bansi' and 'Daebong', respectively. To include different levels of astringency, data for the destructively collected parameters (tannin content, firmness, TSS, and simple sugars) and spectroscopic measurements of the intact fruits were acquired at harvest, on the first day after harvest, and at 3-day intervals afterward until the fruit attain the end of their shelf life (Figure 1). The number of fruits at each measurement day was started with five fruits at the beginning of the storage and decided afterward to 5-10 fruits based on the fruit status.

2.2 VNIR spectral acquisition and analysis

In accordance with Tilahun et al. (2020; Tilahun et al., 2018), each individual intact fruit of the 'Cheongdo-Bansi' and 'Daebong'

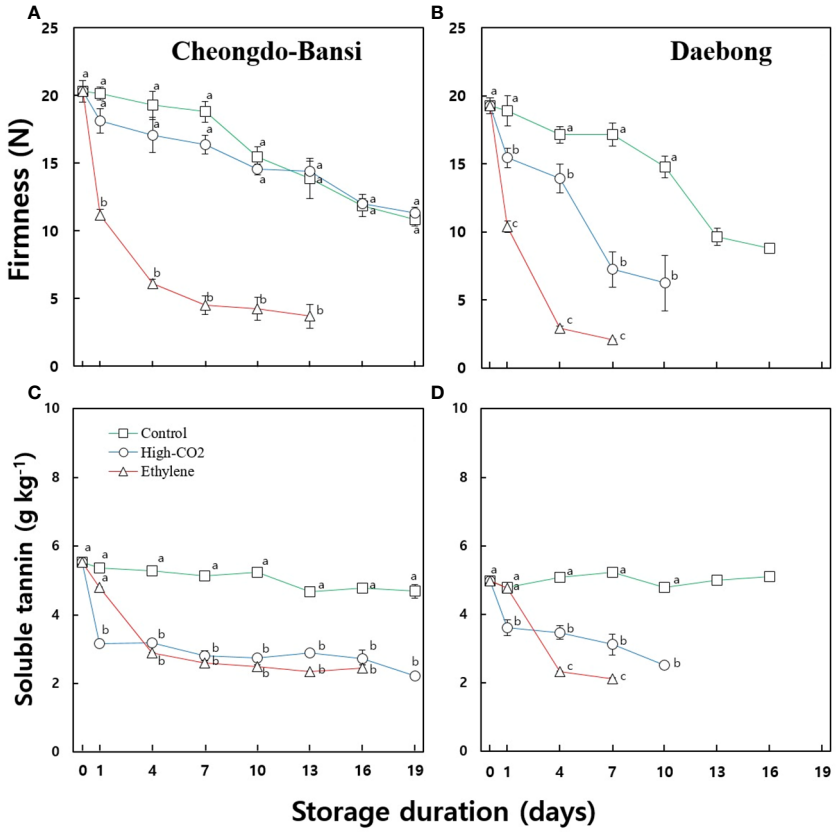


FIGURE 1
Changes in firmness and soluble tannin content of the control and ethylene or high CO_2 treated 'Cheongdo-Bansi' **(A, C)** and 'Daebong' **(B, D)** persimmon fruit during storage at 22°C. Each data point indicates 5-10 fruits. The number of fruit at each measurement day was started with five fruits at the beginning of the storage and decided afterward to 5-10 fruits for each measurement based on the fruit status.

cultivars were scanned in the transmittance mode in the spectral region of 500-1100 nm using three (12 V/100 W) halogen lamps as a source of VNIR light. Fruit holder was used to keeping the fruit right above the detector (Figure 1B). The integration time was set to 100 ms and the measurement was done 12 times from different directions. The fruit was placed on the fruit holder to prevent it from falling, and the fruit holder was rotated above the detector to avoid the interference of scanning by human hand (4 positions on stem-end plane, middle plane, and stylar-end plane (Noypitak et al., 2015) per fruit) to introduce variability within the fruit samples. For each measurement, a total of 3500 data points was captured at 0.2 nm sampling intervals. The VNIR spectrometer was linked to a computer to transfer data. A total of 1440 spectra readings from 120 fruits for 'Cheongdo-Bansi' and 1440 spectra readings from 120 fruits for 'Daebong' were acquired from the intact persimmon fruit throughout the ripening/deastringency period. After removing outliers, a total of 1200 spectra readings (10 readings per fruit) were chosen for analysis from each 'Cheongdo-Bansi' and 'Daebong' (Figure 2). For each cultivar, the spectra readings from half of the total fruit (600 readings from 60 fruits) were utilized for the calibration set, while the other half readings (600 readings from 60 fruits) were used for the prediction set using the leave one sample out procedure to separate the sample sets. The original spectra were transformed by multiplicative scattering correction (MSC), first derivatives, the Hanning window, and standard normal variate (SNV) to remove undesired information and reduce systematic noise. The prediction was based on the lowest predicted residual error sum of squares (PRESS) value, which was used to determine the ideal number of latent variables for the partial least squares regression (PLSR) model. To determine a linear relationship between measured references and spectral data, MATLAB R2012b (Version 8.0.0.783, The Math Works, Inc., Natick, MA, USA) was used to conduct PLSR regression analysis. RMSEC (root mean square of standard error in calibration), RMSEP (root mean square of standard error in prediction), coefficient of determination for calibration (Rc^2) and prediction (Rp^2) were used to evaluate the performance of the developed PLSR models. A predictive model with higher Rp^2, small bias values and lower RMSEP is considered as a reliable prediction model.

2.3 Measurements of fruit quality parameters and analysis

The measurements for firmness, TSS, soluble tannin content, fructose and glucose content were made from each whole fruit according to the methodology implemented in our laboratory and described by Park et al. (2017). A Rheo meter (Sun Scientific Co.

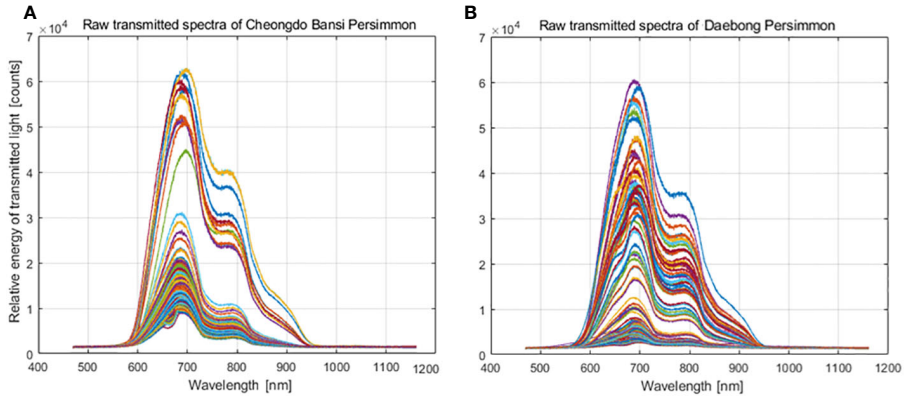


FIGURE 2
Transmittance energy spectra curves obtained from and 'Cheongdo-Bansi' **(A)** and 'Daebong' **(B)** persimmon fruit by using VNIR spectrometer.

Ltd., Tokyo, Japan) with a 10 kgf maximum force of penetration and a 3 mm round, flat-ended stainless-steel probe was utilized to measure the firmness of the intact persimmon fruit with a probe speed of 1 mm/s around equatorial area of each fruit. TSS was measured by utilizing a digital refractometer (Atago Co. Ld., Tokyo, Japan) and 5 g of homogenized persimmon pulp juice from each whole fruit. Glucose and fructose contents were measured in accordance with the method employed by Park et al. (2016); 5 g of each whole fruit's frozen pulp sample was added to 50 mL of distilled water, homogenized, and then the juice was centrifuged (Mega-17R, Hanil Science Industrial, Korea) at 12,578 × g for 10 min and the supernatant was filtered through 0.45 µm membrane filter (Advantec, Tokyo, Japan). The analysis was carried out using HPLC with a RI detector (Waters 410 Differential Refractometer, Waters, MA, USA) and a Sugar-Pak ™ column (6.5 × 300 mm, Waters, USA) with an injection volume of 10 µL. Soluble tannin content was measured with the modified Folin-Dennis method (Park et al., 2017). Samples of 5 g from whole fruit were added directly into a solution of 25 mL of 80% methanol. Then, 6 mL of distilled water was added to 1 mL of filtered supernatant sample solution. The mixture was then vortexed after 0.25 mL of 2 N Folin-Ciacalteau reagent had been added. Saturated Na_2CO_3 (1 mL) and distilled water (1.5 mL) were added after 3 min. Following a 1 h incubation period at 25°C, the absorbance of mixed sample was measured with a spectrometer (Thermo Fisher Scientific, Waltham, MA, USA) at 725 nm, and the results were reported as g kg^{-1} on a fresh weight basis.

To perform PLSR models using the above spectra readings (10 readings per fruit) obtained from different directions of a fruit, measured data for reference parameters (tannin, firmness, SSC, glucose, and fructose) were collected from a total of 240 fruit samples. These samples comprised 120 fruit from each of 'Cheongdo-Bansi' and 'Daebong', with 8 replications per fruit sample, and the mean value of each 4 replicates was used as the fifth value for each parameter to make 10 replicates per fruit to get a one to one fit with the spectra readings.

To examine the effectiveness of multivariate regression models to estimate tannin content (astringency), the values of the collected parameters were divided into calibration and prediction sets using the leave one sample out procedure to separate the four sample sets. Fruit quality parameters were collected from both cultivars throughout the storage period, and the data were divided in to 3:1 ratio. 80 fruit samples were used for calibration, and 40 fruit samples were used for prediction. A total of 240 fruit samples ('Cheongdo-Bansi' and 'Daebong', 120 each) were used for the experiment in 8 replications per fruit sample and the mean value was calculated for analysis. The measured reference parameters (tannin, firmness, SSC, glucose and fructose) were organized in excel, where the rows represented the number of samples (the total of 120 averaged value from 120 persimmons for each cultivar), and the columns represented the number of variables (X and Y variables). The X-variables, or predictors, were the values of measured firmness, SSC, glucose and fructose values associated with each sample. The Y-variables, or response, were the measured tannin values associated with each sample. Multivariate PLSR models were developed from calibration data set and the measured reference data of each parameter were compared to the predicted data obtained by PLSR models in both the calibration and prediction data sets. A predictive model with higher Rp^2, lower RMSEP, and higher ratio of prediction to deviation (RPD) is thought to be a good prediction model. RPD is calculated by the ratio of SD to RMSEC/P, where SD is the standard deviation of the observed parameters. If the RPD value is less than 1.5, the calibration is not usable. When the RPD is between 1.5 and 2.0, it becomes able to distinguish between high and low values, but when it is between 2.0 and 2.5, it becomes possible to make approximate quantitative predictions (Cortés et al., 2016).

3 Results

31 Firmness and soluble tannin content of persimmon during ripening/deastringency

Figure 1 shows the changes in firmness and soluble tannin contents of the control and ethylene or high CO_2 treated

10.3389/fpls.2023.1260644

'Cheongdo-Bansi' and 'Daebong' persimmon fruit during storage at 22°C. The firmness and soluble tannin data showed significant differences between the treatments starting from the first day. Ethylene treated fruits ripened faster, became softer, and acquired a jelly-like consistency which reduced their storage life up to 13 and 7 d for 'Cheongdo-Bansi' and 'Daebong', respectively, compared to the controls that reached 19 and 16 d for 'Cheongdo-Bansi' and 'Daebong', respectively. Conversely, high CO_2 treatment hastened deastringency and maintained firmness and stored up to 19 and 10 d for 'Cheongdo-Bansi' and 'Daebong', respectively. Ethylene treatment hastens softening in both cultivars while high CO_2 maintained firmness and prolonged the storability of both cultivars, and its effect was distinctly higher in the case of 'Cheongdo-Bansi' (Figure 1).

32 VNIR spectra vs. reference analysis

In this study, transmittance energy spectra of 'Cheongdo-Bansi' and 'Daebong' intact astringent persimmon fruit were recorded by VNIR spectrometer (Figures 3A, B) in the wavelength of 500-1000 nm as shown in Figures 2A, B. As indicated in Figure 2, differences were observed in the raw transmitted spectra characteristics of the two cultivars. More scattered spectra were observed in 'Cheongdo-Bansi' than 'Daebong'.

In addition to the PLSR models for the estimation of tannin contents to determine astringency levels of the two persimmon cultivars, PLSR models were also developed to predict postharvest quality parameters such as firmness, TSS, and simple sugars (glucose and fructose). Table 1 shows the essential data for VNIR

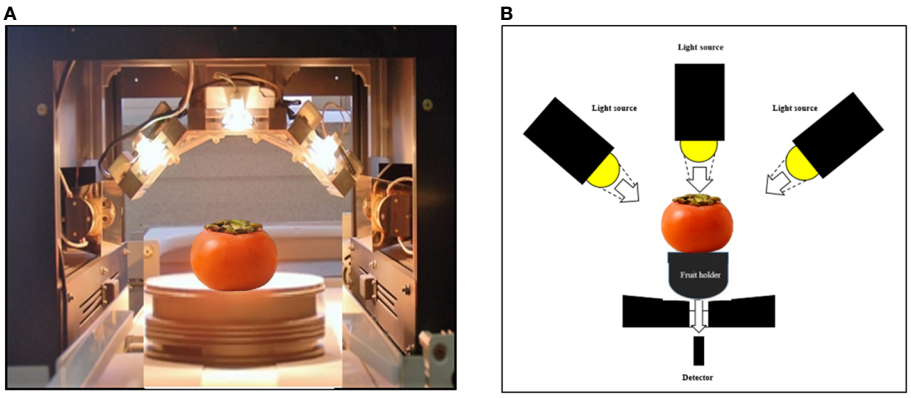


FIGURE 3
VNIR spectrometer **(A)** and measurement system **(B)** during transmittance spectra measurement of intact persimmon fruits.

TABLE 1 Firmness, TSS, simple sugars and soluble tannin content data used for VNIR modeling and multiple regressions.

Cultivars	Parameters	Treatments	Storage duration (days)							
			0	1	4	7	10	13	16	19
Cheongdo-Bansi	Firmness (N)	Control	20.29 ± 0.79	20.14 ± 0.47	19.27 ± 1.02	18.79 ± 0.79	15.48 ± 0.72	13.88 ± 1.49	11.87 ± 0.84	10.83 ± 0.45
		High-CO_2	20.29 ± 0.79	18.12 ± 0.90	17.08 ± 1.29	16.38 ± 0.68	14.54 ± 0.42	14.41 ± 0.75	12.02 ± 0.33	11.34 ± 0.39
		Ethylene	20.29 ± 0.79	11.18 ± 0.39	6.10 ± 0.35	4.5 ± 0.69	4.24 ± 0.87	3.68 ± 0.87	–	–
	TSS (%)	Control	16.93 ± 1.12	17.15 ± 0.79	17.48 ± 0.45	19.41 ± 0.37	19.43 ± 0.24	18.55 ± 0.34	20.04 ± 0.42	19.39 ± 0.56
		High-CO_2	16.93 ± 1.12	15.25 ± 0.43	15.02 ± 0.15	15.83 ± 0.59	16.69 ± 0.18	16.31 ± 0.34	17.10 ± 0.51	18.30 ± 0.70
		Ethylene	16.93 ± 1.12	15.55 ± 0.29	16.90 ± 0.32	16.98 ± 0.55	17.96 ± 0.36	20.11 ± 0.86	21.08 ± 0.63	–
	Glucose (mg kg^{-1})	Control	5822.23	6269.41	6030.09	7558.36	7833.15	7969.14	8618.81	8760.82
		High-CO_2	5822.23	5202.74	5585.59	6070.89	6297.67	6564.06	6630.25	6404.56

(Continued)

10.3389/fpls.2023.1260644

TABLE 1 Continued

Cultivars	Parameters	Treatments	Storage duration (days)							
			0	1	4	7	10	13	16	19
		Ethylene	5822.23	5808.23	6048.38	6044.46	6534.75	6678.41	6448.98	–
	Fructose (mg kg^{-1})	Control	5865.22	6178.98	5891.16	7146.67	7328.13	7461.81	7944.08	7986.70
		High-CO_2	5865.22	4502.56	4751.30	5146.50	5366.50	5523.44	5357.71	5139.43
		Ethylene	5865.22	5300.90	5506.17	5412.71	6110.06	6303.45	5996.22	–
	Soluble tannin (g kg^{-1})	Control	4.68 ± 0.09	4.69 ± 0.03	5.53 ± 0.01	5.23 ± 0.01	4.78 ± 0.07	5.28 ± 0.09	5.14 ± 0.01	5.37 ± 0.21
		High-CO_2	4.68 ± 0.09	3.16 ± 0.07	2.75 ± 0.05	1.80 ± 0.15	3.19 ± 0.01	2.89 ± 0.04	2.72 ± 0.26	2.21 ± 0.01
		Ethylene	4.68 ± 0.09	4.80 ± 0.01	5.17 ± 0.00	4.88 ± 0.01	4.59 ± 0.04	3.82 ± 0.01	3.53 ± 0.09	–
Daebong	Firmness (N)	Control	19.26 ± 0.59	18.89 ± 1.11	17.13 ± 0.60	17.15 ± 0.86	14.79 ± 0.80	9.65 ± 0.62	8.79 ± 0.17	–
		High-CO_2	19.26 ± 0.59	15.44 ± 0.71	13.90 ± 1.06	7.24 ± 1.31	6.23 ± 2.04	–	–	–
		Ethylene	19.26 ± 0.59	10.37 ± 0.41	2.92 ± 0.17	2.05 ± 0.06	–	–	–	–
	TSS (%)	Control	17.48 ± 0.73	18.60 ± 0.41	18.25 ± 0.40	19.76 ± 0.59	18.61 ± 2.20	17.31 ± 0.89	15.20 ± 1.01	–
		High-CO_2	17.48 ± 0.73	15.69 ± 0.28	14.65 ± 0.41	16.40 ± 0.39	15.89 ± 0.45	–	–	–
		Ethylene	17.48 ± 0.73	15.48 ± 0.59	14.81 ± 0.57	16.03 ± 0.33	–	–	–	–
	Glucose (mg kg^{-1})	Control	4933.13	5108.14	5006.48	5371.05	5858.40	5499.64	5861.58	–
		High-CO_2	4933.13	4979.20	4224.52	5210.59	5369.60	–	–	–
		Ethylene	4933.13	4523.83	5006.09	5515.52	–	–	–	–
	Fructose (mg kg^{-1})	Control	4032.12	4162.41	4097.25	4353.54	4727.80	4453.36	4751.39	–
		High-CO_2	4032.12	3788.34	3242.84	3817.81	4010.23	–	–	–
		Ethylene	4032.12	3558.48	3919.92	4174.34	–	–	–	–
	Soluble tannin (g kg^{-1})	Control	4.97 ± 0.03	4.79 ± 0.02	5.09 ± 0.02	5.23 ± 0.02	4.79 ± 0.07	5.00 ± 0.03	5.10 ± 0.10	–
		High-CO_2	4.97 ± 0.03	3.60 ± 0.23	3.46 ± 0.20	3.12 ± 0.31	2.52 ± 0.02	–	–	–
		Ethylene	4.97 ± 0.03	4.76 ± 0.05	2.32 ± 0.01	2.12 ± 0.03	–	–	–	–

Ethylene treated fruits ripened faster, became softer, and acquired a jelly-like consistency which reduced their storage life up to 13 and 7 d for 'Cheongdo-Bansi' and 'Daebong', respectively. The number of fruits at each measurement day was started with five fruits at the beginning of the storage and decided afterward to 5-10 fruits for each measurement based on the fruit status.

modeling and multiple regressions. Promising results were recorded for both cultivars with higher predictive models in 'Daebong' than 'Cheongdo-Bansi'. In 'Cheongdo-Bansi', Rc^2 and RMSEC for measured vs. VNIR values of tannin in the calibration set were 0.81 and 0.83 g kg^{-1}, respectively. Similarly, Rp^2 and RMSEP for measured vs. VNIR values of tannin in the prediction set were 0.75 and 0.52 g kg^{-1}, respectively (Figures 4A, B). On the other hand, in 'Daebong', Rc^2 and RMSEC for measured vs. VNIR values of tannin in the calibration set were 0.96 and 0.21 g kg^{-1}, respectively, while Rp^2 and RMSEP for measured vs. VNIR values of tannin in the prediction set were 0.93 and 0.27 g kg^{-1}, respectively (Figures 4C, D).

The same trends of higher predictive models were also observed in 'Daebong' than 'Cheongdo-Bansi' for measured vs. VNIR values of firmness, TSS and simple sugars. In case of firmness, Rc^2 and RMSEC were 0.94 and 1.31 N, and 0.93 and 1.35 N in 'Daebong' and 'Cheongdo-Bansi', respectively (Figures 5A, C). Correspondingly, Rp^2 and RMSEP were 0.89 and 1.83 N, and 0.91 and 1.49 N for 'Daebong'

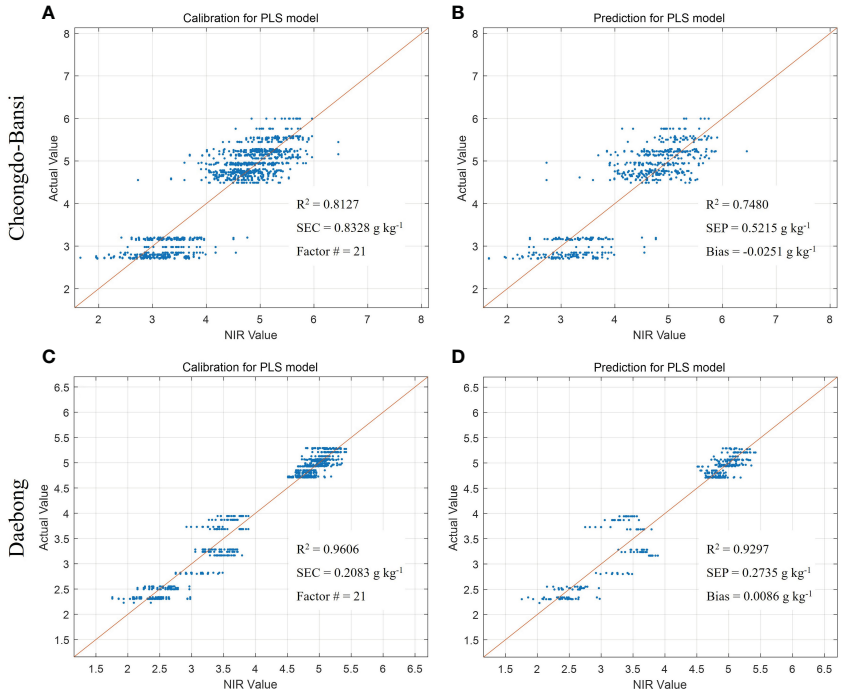


FIGURE 4
Measured vs. predicted values of tannin content (g kg^{-1}) in 'Cheongdo-Bansi' for calibration **(A)** and prediction **(B)**, and 'Daebong' for calibration **(C)** and prediction **(D)** sets with PLS models. The scatter plot depicts the prediction accuracy of the model. The x axis represents the predicted value of tannin content and the y axis represents the measured value by the PLS models.

and 'Cheongdo-Bansi', respectively (Figures 5B, D). Rc^2 and RMSEC for measured vs. VNIR values of TSS were 0.95 and 0.51%, and 0.86 and 0.83%, whereas Rp^2 and RMSEP were 0.93 and 0.55%, and 0.83 and 0.91% in 'Daebong' and 'Cheongdo-Bansi', respectively (Figures 6A–D). Regarding the simple sugars (glucose and fructose), higher predictive models of 0.96 and 0.02 mg kg^{-1}, and 0.96 and 0.02 mg kg^{-1} for Rc^2 and RMSEC, and Rp^2 and RMSEP, respectively, were observed in 'Daebong' (Figures 7C, D). In 'Cheongdo-Bansi', Rc^2 and RMSEC, and Rp^2 and RMSEP were 0.79 and 0.09 mg kg^{-1}, and 0.75 and 0.10 mg kg^{-1}, respectively (Figures 7A, B).

33 Multivariate PLSR models using the reference data

Table 2 shows the means and ranges of reference (measured) tannin content in the calibration and prediction data sets that acquired by the destructive analysis. Meanwhile, tannin content that estimated by multivariate PLSR model using firmness, glucose and fructose in the calibration and prediction data sets are also presented in Table 2. For 'Cheongdo-Bansi', Rc^2, RMSEC and RPD values of the calibration data set were 0.83, 0.27 g kg^{-1} and 0.36, respectively. In the prediction data set, the corresponding values were 0.79, 0.42 g kg^{-1} and 0.22, respectively for Rp^2, RMSEP and RPD (Table 2). Similarly, Rc^2, RMSEC and RPD values of the calibration data set were 0.79, 0.50 g kg^{-1} and 0.38, respectively for 'Daebong'. The corresponding values were 0.84, 0.53 g kg^{-1} and 0.43, respectively for Rp^2, RMSEP and RPD in the prediction data set (Table 2).

Following the predictive analysis in multiple regression, firmness, glucose and fructose values were found to have high predictive *p*-values in the prediction of tannin from the measured reference postharvest quality parameters. The following equations were found to be the best equations.

$$Tannin\ (g\ kg^{-1}) = 3.36 - 0.06\ (firmness) - 19.67(fructose) + 24.54(glucose) - \text{'Cheongdo} - \text{Bansi'}$$

$$Tannin\ (g\ kg^{-1}) = 2.1 + 0.06\ (firmness) - 35.21(fructose) + 47.74(glucose) - \text{'Daebong'}$$

The calibration and prediction set with multivariate PLSR models had shown encouraging results to utilize the models based on the measured reference vs. predicted scores of both cultivars. For the prediction data set, a multivariate PLSR model had the highest coefficient of correlation (0.84) for 'Daebong' and (0.79) for 'Cheongdo-Bansi' (Table 2; Figure 8).

4 Discussion

Rapid ripening, deastringency, and softening in ethylene treated persimmon fruit could be due to rapid expression of ripening-related genes (Park et al., 2017; Park et al., 2019). On the other hand, the deastringency of firm persimmon fruit by high CO_2 treatment could be due to the accumulation of acetaldehyde in the fruit by anaerobic respiration, and the soluble tannins become insoluble as

they react with the acetaldehyde (Cortés et al., 2017). Firmer fruit in high CO_2 treated fruit can be explained by the reduction of respiration rate which in turn inhibit the effect of internal ripening hormone, ethylene (Tilahun et al., 2022).

The introduction of environmentally friendly nondestructive technology like VNIR spectroscopy, which has achieved widespread recognition for assessing food quality, is necessary to meet the present demand for high-quality products (Tilahun et al., 2020). More scattered spectra in 'Cheongdo-Bansi' than 'Daebong' could be attributed to the relatively more sample variation in 'Cheongdo-Bansi' during the extended storage period up to 16, 19, and 19 days for ethylene, high CO_2, and control, respectively. In contrast, 'Daebong' had a shorter storage period of only 7, 10, and 16 days for ethylene, high CO_2, and control, respectively.

From the results of this study, the performance of PLSR models for the prediction of tannin content in intact persimmon fruit was cultivar dependent. Noypitak et al. (2015) evaluated tannin content in high CO_2 treated and control intact 'Xichu' persimmon fruits using NIR and reported PLSR models with 0.94 and 0.95 Rp^2 in transmittance and interactance modes, respectively. They suggested to use reflected light than transmitted light due to the variation of soluble tannin content in the flesh close to the skin and at the core. Cortés et al. (2017) also reported PLSR models using the data obtained from high CO_2 treated intact fruits of 'Rojo Brillante' persimmon at six measurement points in reflectance mode with Rp^2 of 0.90 and 0.91 with all and selected wavelengths, respectively. Nevertheless, our results of this study revealed the possible application of VNIR spectra in transmittance mode to predict tannin content in intact persimmon fruit with higher predictive models in 'Daebong' than 'Cheongdo-Bansi'. Previous studies by Tilahun et al. (2020, Tilahun et al., 2018) on tomatoes and potatoes also support the use of spectra in transmittance mode to predict lycopene, β-carotene, and glycoalkaloids. Moreover, the novelty of this study lies in its incorporation of both deastringency and ripening treatments, encompassing persimmon fruits exhibiting varying degrees of astringency and firmness.

Munera et al. (2017b) reported the potential of hyperspectral imaging to predict firmness with Rp^2 of 0.80 in 'Rojo Brillante' persimmon fruit. Cortés et al. (2016) also predicted internal quality (combination of TSS, firmness and flesh color) of mango with VNIR reflection spectroscopy and reported Rp^2 between 0.83–0.88 using full spectral range. Similarly, Ar et al. (2019) demonstrated the possibility of using NIR spectroscopy to predict TSS and firmness with Rp^2 of 0.86 and 0.94, respectively, in astringent 'Rendeu' persimmon fruit, while there was low accuracy in predicting vitamin C and total acid due to their low contents in persimmon.

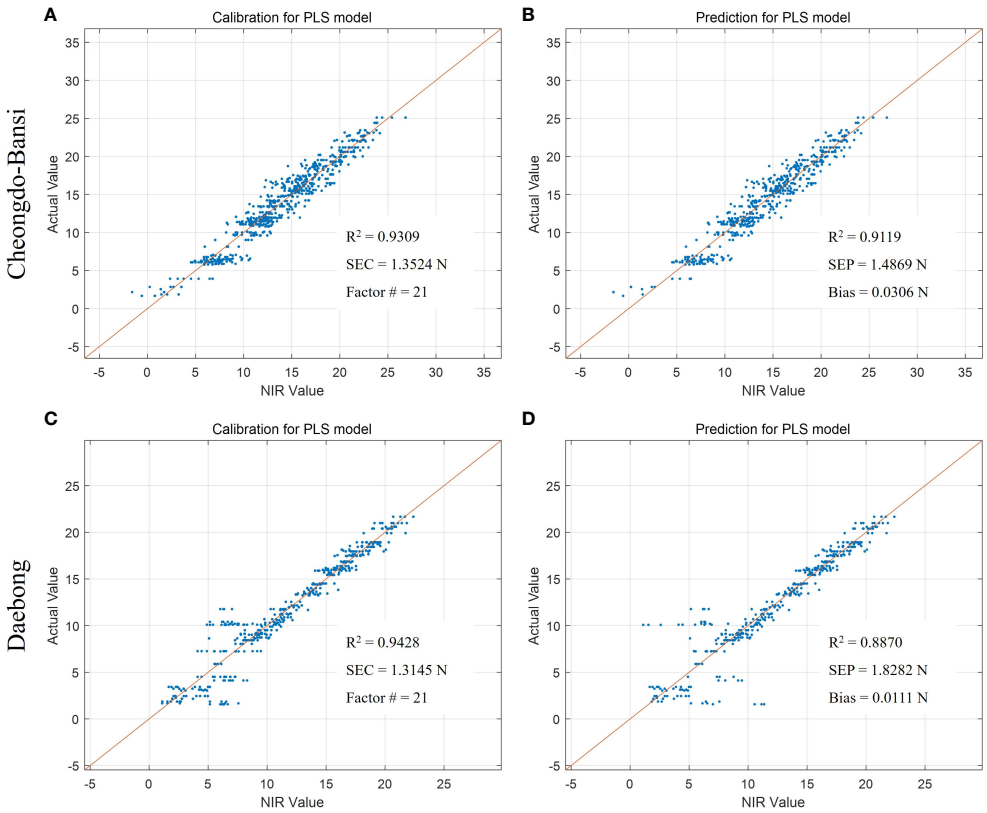


FIGURE 5
Measured vs. predicted values of firmness (N) in 'Cheongdo-Bansi' for calibration **(A)** and prediction **(B)**, and 'Daebong' for calibration **(C)** and prediction **(D)** sets with PLS models. The scatter plot depicts the prediction accuracy of the model. The x axis represents the predicted values of firmness and the y axis represents the measured value by the PLS models.

It is important to note that, in addition to estimating astringency, the developed PLSR models in the current work can be used as the better nondestructive tools for the assessment of the firmness and TSS in both cultivars. Similar to our present study, Liu et al. (2006) reported best models for the prediction of simple sugars in intact apples using Fourier transform near-infrared (FT-NIR) spectroscopy. Taken together, the feasibility of using VNIR to predict all dependent variables (quality parameters and astringency level in terms of tannin content) of persimmon fruit were indicated by lower RMSEC/P values and higher Rp^2 between 0.89-0.96 and 0.75-0.91 for 'Daebong' and 'Cheongdo-Bansi', respectively. The wide range of NIR values in the developed PLSR models could be due to ten spectra readings obtained from different directions from one fruit, whereas eight reference measured values were collected per fruit. In addition, the variation in the nature of the astringency treatments (control, high CO_2, and ethylene) has led to variations in fruit characteristics. Notably, tannin content decreased in both high CO_2 and ethylene treatments, contributing to a narrower range of actual tannin content values.

In our previous works, the multivariate PLSR models were developed to predict lycopene and β-carotene in tomatoes and glycoalkaloids in potatoes from Hunter's color values (Tilahun et al., 2018; Tilahun et al., 2020). Measurements of postharvest quality parameters such as color values, firmness, TSS, and simple sugars (glucose and fructose) were taken during the experiment. However, in this study, the PLSR models for color values and TSS in the calibration data set had lower R^2 and the *p*-values were higher than 0.15 for both 'Cheongdo-Bansi' and 'Daebong'. Hence, color values and TSS were not included in multivariate PLSR model development. Instead, we included simple sugars (glucose and fructose) data for PLSR model development. The above indicated multivariate PLSR models could not be claimed as nondestructive estimation method as they utilize the destructively acquired data for model development. However, astringency levels can be estimated from firmness and simple sugars without the extra analysis of tannin content. This in turn, reduce time, cost of skilled man power and solvents, and does not require expensive high-performance equipment. As the present study incorporated only two cultivars, further studies are needed on various cultivars to develop more robust multivariate PLSR models.

5 Conclusions

The present study indicates attempts to predict tannin content and quality parameters in intact persimmon fruit with chemical-free, fast and cheap VNIR spectra. Multivariate PLSR models were also developed from the reference measured parameters including

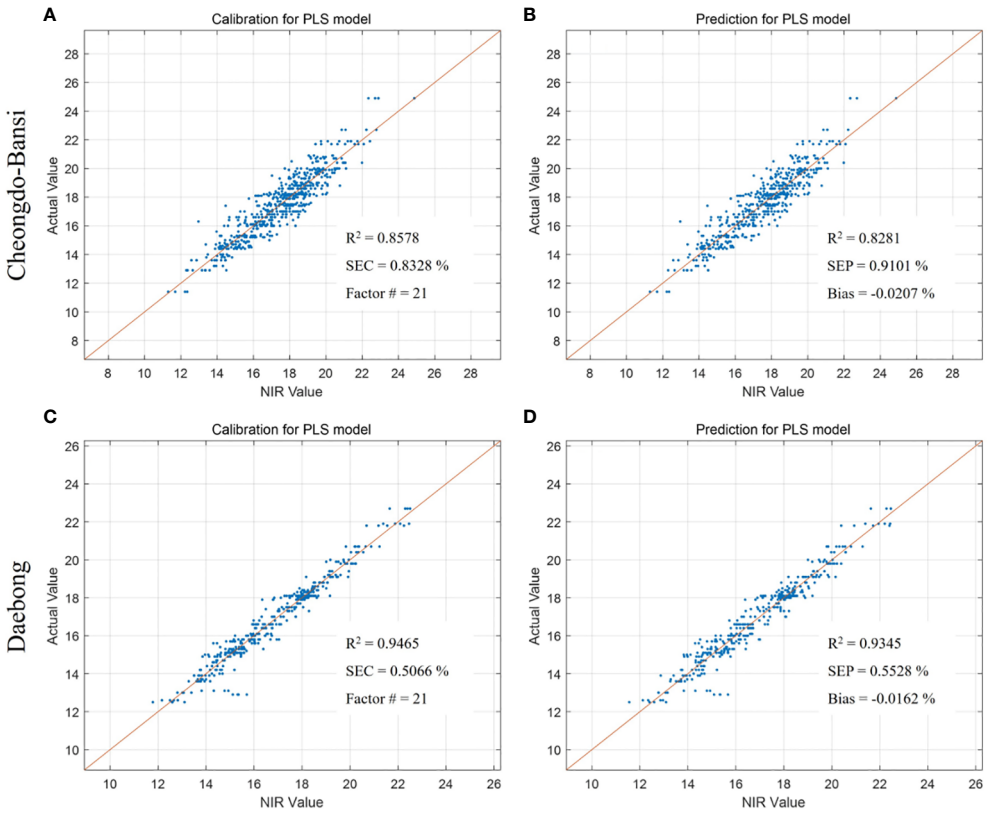


FIGURE 6
Measured vs. predicted values of TSS (%) in 'Cheongdo-Bansi' for calibration **(A)** and prediction **(B)**, and 'Daebong' for calibration **(C)** and prediction **(D)** sets with PLS models. The scatter plot depicts the prediction accuracy of the model. The x axis represents the predicted value of TSS and the y axis represents the measured value by the PLS models.

10.3389/fpls.2023.1260644

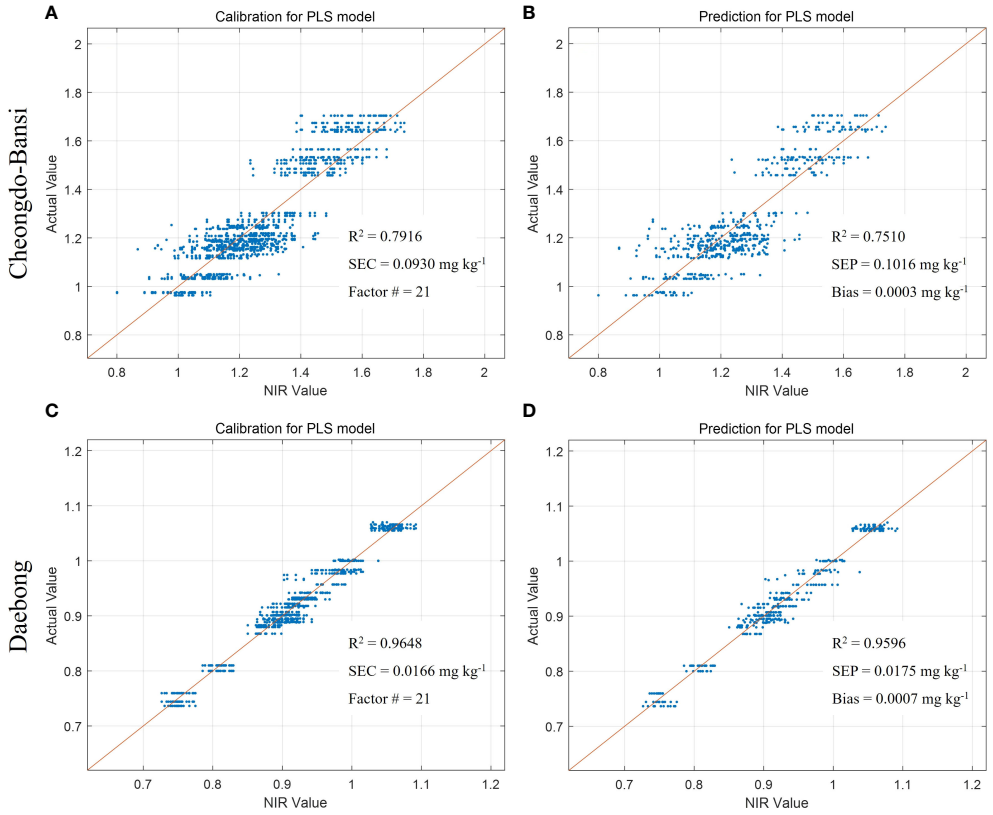


FIGURE 7
Measured vs. predicted values of simple sugars (fructose and glucose (mg kg^{-1})) in 'Cheongdo-Bansi' for calibration **(A)** and prediction **(B)**, and 'Daebong' for calibration **(C)** and prediction **(D)** sets with PLS models. The scatter plot depicts the prediction accuracy of the model. The x axis represents the predicted value of simple sugars and the y axis represents the measured value by the PLS models.

TABLE 2 Statistics for multivariate calibration and prediction of tannin content in 'Cheongdo-Bansi' and 'Daebong' persimmon fruit.

Cultivar	Set	Parameters	Fruit number	Total samples	Mean	Range	SD	Rc/ p^2	RMSEC/ P	RPD
Cheongdo-Bansi	Calibration	Reference tannin content	80	640	4.49	2.80-5.44	0.84			
		Multivariate			4.38	2.35-5.45	0.78	0.83	0.36	2.3
	Prediction	Reference tannin content	40	320	4.19	2.73-5.44	1.07			
		Multivariate			4.24	2.58-5.44	0.94	0.79	0.49	2.2
Daebong	Calibration	Reference tannin content	80	640	4.10	2.30-5.29	1.05			
		Multivariate			3.95	2.20-5.40	0.96	0.79	0.50	2.1
	Prediction	Reference tannin content	40	320	4.03	2.30-5.29	1.20			
		Multivariate			4.20	2.42-5.37	0.97	0.84	0.53	2.3

SD, standard deviation; RMSEC, root mean square error of calibration; RMSEP, root mean square error of prediction; RPD, residual prediction deviation (SD/RMSEC/P); Rc^2, coefficient of determination in calibration; Rp^2, coefficient of determination in prediction data set.

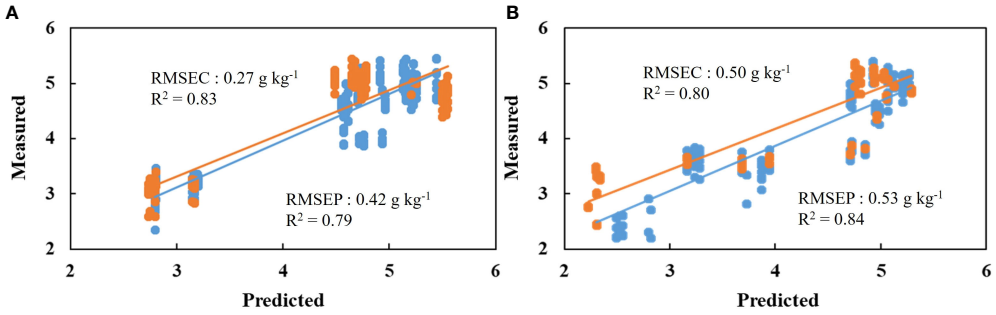


FIGURE 8
Measured vs. predicted scores of tannin content (g kg^{-1}) for 'Cheongdo-Bansi' **(A)** and 'Daebong' **(B)** in the calibration (blue) and prediction (red) sets with multivariate PLS models using firmness and simple sugars values. The scatter plot depicts the prediction accuracy of the model. The x axis represents the predicted value of tannin content and the y axis represents the measured value by the multivariate PLS models.

firmness, TSS, and simple sugars. Prediction of tannin content, firmness, TSS, and simple sugars was promising in both cultivars, and relatively better predictive models were developed in 'Daebong' than 'Cheongdo-Bansi' with both VNIR and multivariate-based techniques. Our models could be promising alternative tools to the costly and time-consuming destructive analysis. The developed models could benefit both the industry and consumers through their use in the agricultural processing and distribution centers to sort fruits on a conveyor belt at different levels of astringency and ripening stages with a VNIR spectrometer. In addition, astringency levels can be estimated from firmness and simple sugars by the developed multivariate PLSR models without the extra analysis of tannin content. Further investigation on different cultivars at different levels of astringency and softening to evaluate tannin content and ripening quality of intact persimmon fruit could help to develop more robust models.

Data availability statement

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

Author contributions

MB: Conceptualization, Formal analysis, Writing – original draft, Writing – review & editing, Data curation, Investigation, Methodology. HC: Formal analysis, Investigation, Methodology, Writing – review & editing. IH: Formal analysis, Methodology, Writing – review & editing, Software, Validation. ST: Formal analysis, Writing – review & editing, Conceptualization, Writing – original draft. CJ: Conceptualization, Funding acquisition, Project administration, Supervision, Writing – review & editing.

Funding

The author(s) declare financial support was received for the research, authorship, and/or publication of this article. This study was supported by the National Research Foundation of Korea (NRF) grant funded by the Korea government (MIST) (RS-2023-00240947) and BK21 FOUR project.

Conflict of interest

Author IH is employed by Frusen 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.

References

Akagi, T., Ikegami, A., Tsujimoto, T., Kobayashi, S., Sato, A., Kono, A., et al. (2009). DkMyb4 is a Myb transcription factor involved in proanthocyanidin biosynthesis in persimmon fruit. *Plant Physiol.* 151, 2028–2045. doi: 10.1104/pp.109.146985

Ar, N. H., Purwanto, Y. A., and Budiastra, I. W. (2019). "Prediction of soluble solid content, vitamin C, total acid and firmness in astringent persimmon (*Diospyros kaki* L.) cv. Rendeu using NIR spectroscopy," in *IOP Conference Series: Materials Science and*

Engineering. 12086 (IOP Publishing). The 1st International Conference on Mechanical Electronic and Biosystem Engineering (MEBSE 2018) Bogor, Indonesia

Arnal, L., and Del Río, M. A. (2003). Removing astringency by carbon dioxide and nitrogen-enriched atmospheres in persimmon fruit cv. "Rojo brillante". *J. Food Sci.* 68, 1516–1518. doi: 10.1111/j.1365-2621.2003.tb09676.x

Arnal, L., Martínez-Jávega, J. M., Besada, C., and Salvador, A. (2008). "Optimization of the duration of deastringency treatment depending on persimmon maturity," in *III international conference postharvest unlimited*, International Society for Horticultural Science, ISHS Acta Horticulturae: Berlin, Germany Vol. 858. 69–74.

Ben-Arie, R., and Sonego, L. (1993). Temperature affects astringency removal and recurrence in persimmon. *J. Food Sci.* 58, 1397–1400. doi: 10.1111/j.1365-2621.1993.tb06191.x

Cortés, V., Ortiz, C., Aleixos, N., Blasco, J., Cubero, S., and Talens, P. (2016). A new internal quality index for mango and its prediction by external visible and near-infrared reflection spectroscopy. *Postharvest Biol. Technol.* 118, 148–158.

Cortés, V., Rodríguez, A., Blasco, J., Rey, B., Besada, C., Cubero, S., et al. (2017). Prediction of the level of astringency in persimmon using visible and near-infrared spectroscopy. *J. Food Eng.* 204, 27–37. doi: 10.1016/j.jfoodeng.2017.02.017

Das, P. R., and Eun, J. B. (2021). Removal of astringency in persimmon fruits (*Diospyros kaki*) subjected to different freezing temperature treatments. *J. Food Sci. Technol.* 58, 3154–3163. doi: 10.1007/s13197-020-04818-3

FAOSTAT (2021) *Food & Agriculture Organization of the United Nations Statistics Division*. Available at: http://faostat3.fao.org/home/index.html (Accessed 15 February 2023).

Liu, Y., Ying, Y., Yu, H., and Fu, X. (2006). Comparison of the HPLC method and FT-NIR analysis for quantification of glucose, fructose, and sucrose in intact apple fruits. *J. Agric. Food Chem.* 54, 2810–2815. doi: 10.1021/jf052889e

Munera, S., Besada, C., Aleixos, N., Talens, P., Salvador, A., Sun, D. W., et al. (2017a). Non-destructive assessment of the internal quality of intact persimmon using colour and VIS/NIR hyperspectral imaging. *Lwt* 77, 241–248. doi: 10.1016/j.lwt.2016.11.063

Munera, S., Besada, C., Blasco, J., Cubero, S., Salvador, A., Talens, P., et al. (2017b). Astringency assessment of persimmon by hyperspectral imaging. *Postharvest Biol. Technol.* 125, 35–41. doi: 10.1016/j.postharvbio.2016.11.006

Novillo, P., Salvador, A., Llorca, E., Hernando, I., and Besada, C. (2014). Effect of CO2 deastringency treatment on flesh disorders induced by mechanical damage in persimmon. *Biochem. microstructural Stud. Food Chem.* 145, 454–463. doi: 10.1016/j.foodchem.2013.08.054

Noypitak, S., Terdwongworakul, A., Krisanapook, K., and Kasemsumran, S. (2015). Evaluation of astringency and tannin content in xichu persimmons using near infrared spectroscopy. *Int. J. Food Prop.* 18, 1014–1028. doi: 10.1080/10942912.2014.884577

Ortiz, G. I., Sugaya, S., Sekozawa, Y., Ito, H., Wada, K., and Gemma, H. (2005). Efficacy of 1-Methylcyclopropene (1-MCP) in Prolonging the Shelf-life of ‘ Rendaiji’ Persimmon fruits previously subjected to astringency removal treatment. *J. Japanese Soc Hortic. Sci.* 74, 248–254. doi: 10.2503/jjshs.74.248

Parfitt, D. E., Yonemori, K., Honsho, C., Nozaka, M., Kanzaki, S., Sato, A., et al. (2015). Relationships among asian persimmon cultivars, astringent and non-astringent types. *Real-Time Syst.* 51, 1–9. doi: 10.1007/s11295-015-0848-z

Park, D. S., Tilahun, S., Heo, J. Y., and Jeong, C. S. (2017). Quality and expression of ethylene response genes of 'Daebong' persimmon fruit during ripening at different temperatures. *Postharvest Biol. Technol.* 133, 57–63. doi: 10.1016/j.postharvbio.2017.06.011

Park, D.-S., Tilahun, S., Hyun, J.-Y., Kwon, H.-S., and Jeong, C.-S. (2016). Effect of ripening conditions on quality of winter squash "bochang". *Korean J. Food Sci. Technol.* 48(2), 142–146. doi: 10.9721/kjfst.2016.48.2.142

Park, D. S., Tilahun, S., Park, K. C., Choi, I. Y., and Jeong, C. S. (2019). Transcriptome analysis of astringent 'Cheongdo-Bansi' persimmon fruit treated with ethylene for removal of astringency. *Postharvest Biol. Technol.* 150, 52–59. doi: 10.1016/j.postharvbio.2018.12.014

Pasquini, C. (2003). Near infrared spectroscopy: Fundamentals, practical aspects and analytical applications. *J. Braz. Chem. Soc* 14, 198–219. doi: 10.1590/S0103-50532003000200006

Salvador, A., Arnal, L., Besada, C., Larrea, V., Quiles, A., and Pérez-Munuera, I. (2007). Physiological and structural changes during ripening and deastringency treatment of persimmon fruit cv. "Rojo Brillante". *Postharvest Biol. Technol* 46, 181–188. doi: 10.1016/j.postharvbio.2007.05.003

Son, J. R., Lee, K. J., Kang, S., Kim, G., Yang, G. M., Mo, C. Y., et al. (2009). Development of prediction models for nondestructive measurement of sugar content in sweet persimmon. *J. Biosyst. Eng.* 34, 197–203.

Tilahun, S., An, H. S., Hwang, I. G., Choi, J. H., Baek, M. W., Choi, H. R., et al. (2020). Prediction of α -solanine and α -chaconine in potato tubers from hunter color values and VIS/NIR spectra. *J. Food Qual.* doi: 10.1155/2020/8884219

Tilahun, S., Jeong, M. J., Choi, H. R., Baek, M. W., Hong, J. S., and Jeong, C. S. (2022). Prestorage high CO_2 and 1-MCP treatment reduce chilling injury, prolong storability, and maintain sensory qualities and antioxidant activities of "madoka" peach fruit. *Front. Nutr.* 9. doi: 10.3389/fnut.2022.903352

Tilahun, S., Park, D. S., Seo, M. H., Hwang, I. G., Kim, S. H., Choi, H. R., et al. (2018). Prediction of lycopene and β-carotene in tomatoes by portable chroma-meter and VIS/NIR spectra. *Postharvest Biol. Technol.* 136, 50–56. doi: 10.1016/j.postharvbio.2017.10.007

Yamada, M., Taira, S., Ohtsuki, M., Sato, A., Iwanami, H., Yakushiji, H., et al. (2002). Varietal differences in the ease of astringency removal by carbon dioxide gas and ethanol vapor treatments among oriental astringent persimmons of Japanese and chinese origin. *Sci. Hortic. (Amsterdam).* 94, 63–72. doi: 10.1016/S0304-4238(01)00367-3

frontiers | Frontiers in Plant Science

TYPE Review
PUBLISHED 10 January 2024
DOI 10.3389/fpls.2023.1324881

OPEN ACCESS

EDITED BY
Lie Deng,
Southwest University, China

REVIEWED BY
Yanru Zhao,
Northwest A&F University, China
Zhiming Guo,
Jiangsu University, China
Sen Nie,
China Agricultural University, China

*CORRESPONDENCE
Han Zhang
zhangfan67@126.com
Wei Tang
wtang906@163.com

RECEIVED 20 October 2023
ACCEPTED 19 December 2023
PUBLISHED 10 January 2024

CITATION
Xia Y, Liu W, Meng J, Hu J, Liu W, Kang J, Luo B, Zhang H and Tang W (2024) Principles, developments, and applications of spatially resolved spectroscopy in agriculture: a review.
Front. Plant Sci. 14:1324881.
doi: 10.3389/fpls.2023.1324881

Principles, developments, and applications of spatially resolved spectroscopy in agriculture: a review

Yu Xia[1], Wenxi Liu[1,2], Jingwu Meng[1], Jinghao Hu[1], Wenbo Liu[1], Jie Kang[1], Bin Luo[2], Han Zhang[2*] and Wei Tang[1*]

[1]School of Electrical and Control Engineering, Shaanxi University of Science & Technology, Xi'an, Shaanxi, China, [2]Intelligent Equipment Research Center, Beijing Academy of Agriculture and Forestry Sciences, Beijing, China

Agriculture is the primary source of human survival, which provides the most basic living and survival conditions for human beings. As living standards continue to improve, people are also paying more attention to the quality and safety of agricultural products. Therefore, the detection of agricultural product quality is very necessary. In the past decades, the spectroscopy technique has been widely used because of its excellent results in agricultural quality detection. However, traditional spectral inspection methods cannot accurately describe the internal information of agricultural products. With the continuous research and development of optical properties, it has been found that the internal quality of an object can be better reflected by separating the properties of light, such as its absorption and scattering properties. In recent years, spatially resolved spectroscopy has been increasingly used in the field of agricultural product inspection due to its simple compositional structure, low-value cost, ease of operation, efficient detection speed, and outstanding ability to obtain information about agricultural products at different depths. It can also separate optical properties based on the transmission equation of optics, which allows for more accurate detection of the internal quality of agricultural products. This review focuses on the principles of spatially resolved spectroscopy, detection equipment, analytical methods, and specific applications in agricultural quality detection. Additionally, the optical properties methods and direct analysis methods of spatially resolved spectroscopy analysis methods are also reported in this paper.

KEYWORDS

spatially resolved spectroscopy, optical properties, quality inspection, agriculture, hyperspectral imaging

1 Introduction

With the improvement of living standards, people have higher and higher requirements for the quality and safety of agricultural products (Rejeb et al., 2022). Nondestructive testing techniques for the quality of agricultural products have also become more and more widespread in recent years (Tian et al., 2023). With the development of optical technology, some efficient and mature optical nondestructive detection techniques have emerged (Mei and Li, 2023; Mohd Ali et al., 2023), such as visible and near-infrared (Vis-NIR) spectroscopy (Guo et al., 2023) and hyperspectral imaging (HSI) (Chen et al., 2021; Tian et al., 2021; Zhang et al., 2022; Zhao et al., 2023), which have been widely used in nondestructive quantities for physical and chemical characterization of agricultural products. These optical inspection techniques can be mainly used to measure the spectral information of agricultural products to obtain the diffuse reflectance (or transmittance) of the samples and then combine this spectral information with existing chemometrics algorithms to establish a prediction model for the quality of agricultural products. Although existing intelligent information processing techniques are more mature, such as deep learning and machine learning (Audu and Aremu, 2021; Dhanya et al., 2022; Ryo, 2022), these methods have been widely developed and can further enhance the ability to detect the quality of agricultural products. Nevertheless, the spectral information which has already been obtained, can be only analyzed by these methods, and if the spectral information obtained is better, then the quality of agricultural products will be more accurately detected. When light enters the surface of an object, a series of optical phenomena such as scattering and absorption will occur, and this optical information is very important for the detection of the quality of the spectrum. The common spectral acquisition methods often produce significant errors and cannot accurately describe the absorption and scattering information of the light. In order to describe more accurately the laws of propagation of light in the organization of an object as well as more specific properties, special studies have been made on optical properties (OP).

When light enters a turbid medium, a series of optical phenomena occur, such as reflection, refraction, absorption, and scattering. Absorption and scattering of light are the most dominant OP of light in biological tissues. The absorption coefficient (μ_a) and the reduced scattering coefficient (μ_s') are specific descriptions of the absorption and scattering properties of light. The μ_a is mainly related to the chemical composition of the biological tissue, while the μ_s' is closely associated with the structural and physical properties of the sample tissue. Conventional optical inspection techniques can only detect the total effect of light absorption and scattering, but it is not easy to measure the specific parameters of these OP accurately. Researchers have made great efforts to distinguish between scattering and light absorption effects in tissues. Currently, indirect measurement techniques for optical parameters, represented by time-resolved (TR) (Cubeddu et al., 2001; Zude et al., 2011; Vanolia et al., 2020), spatial-resolved (SR) (Ma et al., 2021b; Huang et al., 2022), frequency-domain (FD) (Hu et al., 2020a), spatial-frequency domain imaging (SFDI) (Hu et al., 2018; Sun Z. et al., 2021) and integrating sphere (IS), are used by measuring intact or partial tissue via obtaining certain specific parameters (such as diffuse reflectance R, diffuse transmittance T, and collimated transmittance T_c, etc.) of intact tissue or slices and combining them with specific optical transmission models and inversion algorithms, the optical parameters of the sample can be obtained indirectly, and the absorption and scattering properties of tissues from light can be separated or obtained simultaneously, thus the chemical and physical information of sample can be eventually reflected. Compared with other detection techniques towards optical properties, spatially resolved spectroscopy (SRS) is simple, low cost, and is widely used and relatively mature in nondestructive testing of agricultural products.

SRS was initially used in the medical field to determine the absorption and scattering properties of light in blood with two parallel optical fibers (Reynolds et al., 1976). This technique is mainly used to measure the diffuse reflection of light at different distances from the sample surface via a point light source and to calculate the absorption and reduced scattering coefficients of light in biological tissues by combining the diffuse reflection equation of light. It has a banana-shaped transmission path, as shown in Figure 1. As the distance between the light source and the detector increases, the SRS method can detect deeper, which can obtain more information about the interior of the corresponding tissue and facilitate the detection of features inside biological tissues. In summary, SRS is a convenient tool for obtaining spectral information at different locations. Since SRS integrates spatial and spectral information, it can help researchers to explore its correlation with the chemical composition, physical structure and OP of the samples to be measured, and to build corresponding prediction models for the purpose of product quality prediction, which has resulted in a wider application of this technology in more fields. For example, agriculture (Huang et al., 2022), forestry (Ma et al., 2021c), industrial construction (Wang et al., 2022), physical and chemical materials (Bao et al., 2021; Liu et al., 2022), astronomical observation (Bao et al., 2021; Comerford et al., 2022), gas detection (Li et al., 2021), biomedicine (Niwayama and Unno, 2021; De Man et al., 2023) and other fields, providing people with crucial scientific basis and reliable data support. In the field of biomedicine, SRS is widely used in human hemoglobin detection,

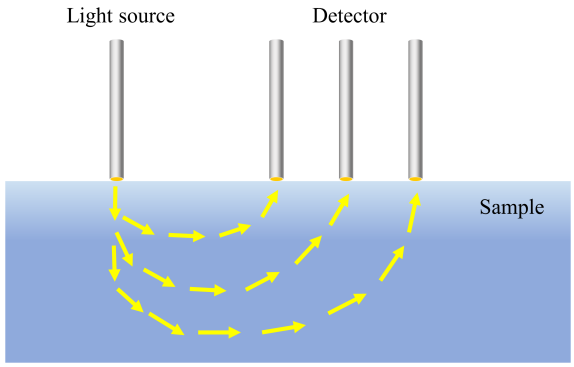


FIGURE 1
Transmission paths of light.

skin pathology detection, and so on. It can help doctors to more accurately identify hemoglobin levels and the human condition so that they can precisely analyze the cause of the patient illness (Zaytsev et al., 2022; Zhang et al., 2023). It can also be used to identify blood species to enhance wildlife protection and preserve national resource information (Zhang et al., 2021). In the field of agriculture, SRS technology is more advantageous than traditional spectral detection technology, and the prediction model established by this technology can improve the prediction ability of the quality for agricultural products, which is currently mainly applied in the quality detection of SSC (Soluble Solids Content), firmness, pH, bruise detection, etc. (Huang et al., 2018b). It can be seen that SRS has a very wide range of utilization in detection with a broad application prospect.

Traditional detection can only obtain the spectrum of a certain place in the sample without gaining more information, and it often collects the total effect of absorption and scattering of light, which may lead to inaccurate prediction results. In contrast, the SRS method can detect spectra at different distances to obtain more spectral information. Moreover, the technique has the advantage of separating the optical properties to analyze the quality of the sample in a targeted manner. Currently, there are many studies based on SRS in the quality inspection of agricultural products, such as the inspection of fruits, meat products, and milk. Since its detection methods establish models that can predict the quality of agricultural products more accurately, it has been widely used in the field of agriculture. There are fewer researchers who have summarized the principles, development, and applications of SRS in agriculture. Therefore, the main objective of this paper is to provide a systematic introduction to different SRS systems and to review the fundamentals, recent developments, and applications of SRS in agricultural quality inspection. In addition, although SRS has been relatively mature in agriculture, it is still faced with many challenges and difficulties presently. The development status and development trend of SRS techniques in agriculture are also reported.

2 Spatially resolved spectral detection systems

With the development of SRS technology, the application field has become more widespread. When using this technique to detect different kinds of samples, people find that some traditional test samples are difficult to meet the needs of different varying detection samples. Not only are there irregularities in the tested models, but modern developments are also demanding faster, more convenient, and more efficient detection configurations or systems, as well as higher detection accuracy and lower device costs. Therefore, researchers are constantly researching and developing more appropriate spatially resolved related detection systems. The following are the existing spatially resolved spectral detection systems at this stage, which mainly include single-fiber, fiber-array, charge-coupled device (CCD) line-scan, hyperspectral line-scan, and multi-channel hyperspectral detection systems.

2.1 Single fiber system

The earliest form of spatially resolved spectral detection was to detect spectral information from different distances by two parallel optical fibers in contact with the object under test (Reynolds et al., 1976). This approach is known as the single fiber system. Only two optical fibers are needed; one is connected to a light source to provide a stable optical signal, and another is connected to a spectrometer to receive the signal. The two fibers follow a certain distance to obtain spatially resolved spectra. This type of format is the simplest, but a large error still exists; it is hard to ensure that the light source-detector distance (SDD) is accurate as well as stable when the fiber is moving, and the two fibers also must be as close as possible to the object under test, so as to avoid the impact of stray light on the quality of the spectral information.

To avoid the impact of manual detection on the experiment, Xia et al. fixed the light source fiber and the detection fiber by a mechanical device (Xia et al., 2007), as shown in Figure 2, which used a 20 W halogen lamp (HL-2000-FHSA-HP, Ocean Optics Inc., Dunedin, USA) as the light source. It is connected to an optical fiber and illuminates the sample surface at an incidence angle of 40°. The detection fiber, connected to the spectrometer, is perpendicular to the sample surface. The position of the detection fiber is moved by a translation device to detect the spectral information at different distances. Both the source and detection fibers have a core diameter of 400 μm, and the closest distance between the two fibers is 1.5 mm to avoid fiber collisions.

For more portability and ease of operation, Ye et al. developed a slidable ring device consisting of a halogen lamp LA-150ue-A (Hayashi Co., Japan), a removable detection fiber ring

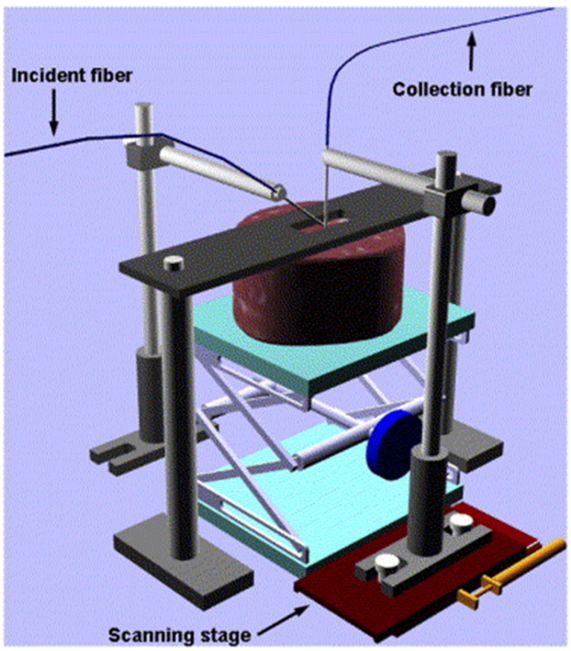


FIGURE 2
Single fiber optic inspection device (Xia et al., 2007).

illuminator, and a Mini-Spectrometer BLACK-Comet-SR100 (StellarNet Inc., USA) (Ye et al., 2021), as shown in Figure 3A. Measuring with the ring illuminator close to the surface of the fruit (Figure 3B). Figure 3C shows a schematic diagram of the ring illuminator. A halogen light source enters the ring illuminator through an optical fiber to form a ring beam, as shown in Figure 3D. As the device inside and outside the machine have the effect of shading to reduce the reflection of light from the sample surface, the spectrum is received only through the small hole in the middle of the signal to reduce the impact of mixed spectral information. The detector and light source are in contact with the sample, and the spectral information is detected by moving the position of the detection fiber in the center of the ring.

The single fiber moving detection form is simple in structure, easy to operate, low cost, and flexible. It can select the optimal detection SDD so that the collected information is more representative. However, this method is easily affected by many factors, such as the accuracy and stability of the moving platform, the strength of the light source fiber and the acquisition fiber fixed, the extent of contact between the measured sample irregularities and the acquisition fiber, all of which can make the system have a significant error. In addition, the single fiber detection form has a high demand on the fiber diameter, which requires the fiber diameter to be as thin as possible so that the light SDD can be closer. Xia et al. reduced the light SDD because of the limitation of the fiber diameter adjusted, thereby adjusting the incident light angle (Xia et al., 2007). Furthermore, the time required for single fiber detection is long. Therefore, the use of a single fiber detection format is not friendly for collecting a large number of samples, and a faster and more efficient detection system needs to be developed.

2.2 Fiber array type system

Due to the significant error of a single fiber optic collection system, the acquisition process of each distance can only follow the sequence to collect, which is time-consuming and laborious, and there will be a phenomenon of missed collection, so the researchers have developed a form of detection based on fiber optic arrays to achieve the simultaneous acquisition of multiple distances. Zhou et al. evaluated the OP of turbid media utilizing a multi-fiber detection format (Zhou et al., 2015), as shown in Figure 4. The system was used to collect spatially resolved diffuse reflections at 633 nm with a light source (HL-2000, Ocean Optics, USA), an illumination fiber, six detection fibers, and a spectrometer (QE65pro, Ocean Optics, USA). All the acquisition fibers are connected to a multiplexer, and the spectral signal is transmitted to the spectrometer through the multiplexer.

Spectral information for every distance can be read by a fiber array device connected to the multiplexer. Nevertheless, the sequential reading of each spectral information needs to be set up, and the setup is complicated with a longer reading time. Nguyen Do Trong et al. investigated a new SRS fiber array device (Nguyen Do Trong et al., 2011), shown in Figure 5, which consists of a halogen light source (AvaLight-DHc, Avantes, Netherlands), an illumination fiber and five detection fibers, a spectrometer, a CCD

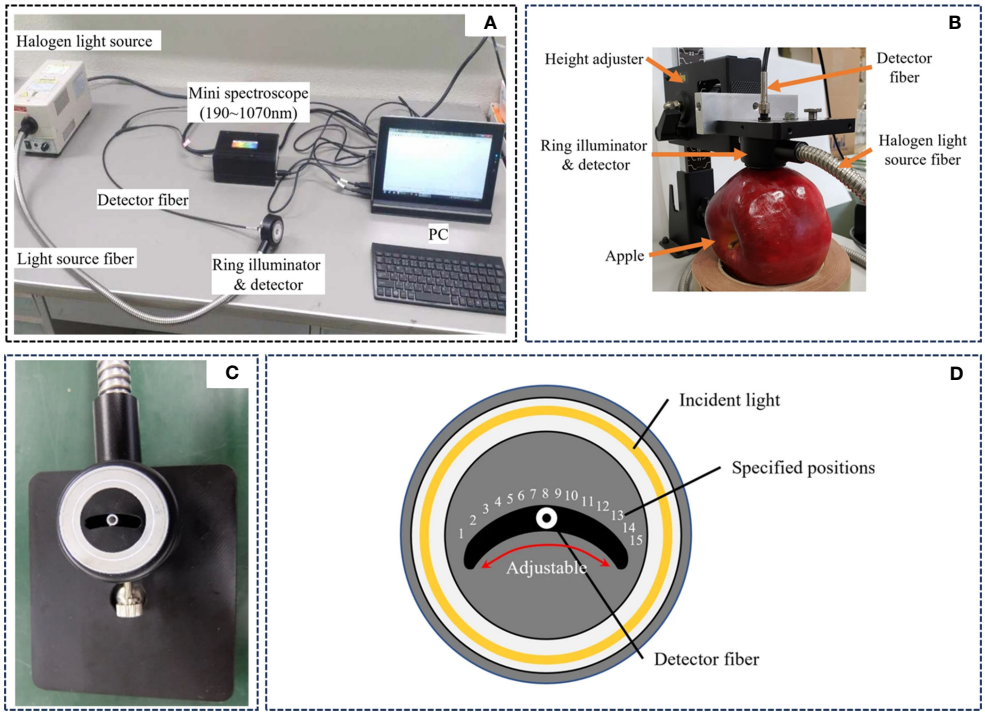


FIGURE 3
Removable probe fiber optic ring device (Ye et al., 2021). **(A)** Slideable ring inspection system. **(B)** Inspection demonstration image. **(C)** Ring illuminator object diagram. **(D)** Ring illuminator schematic diagram.

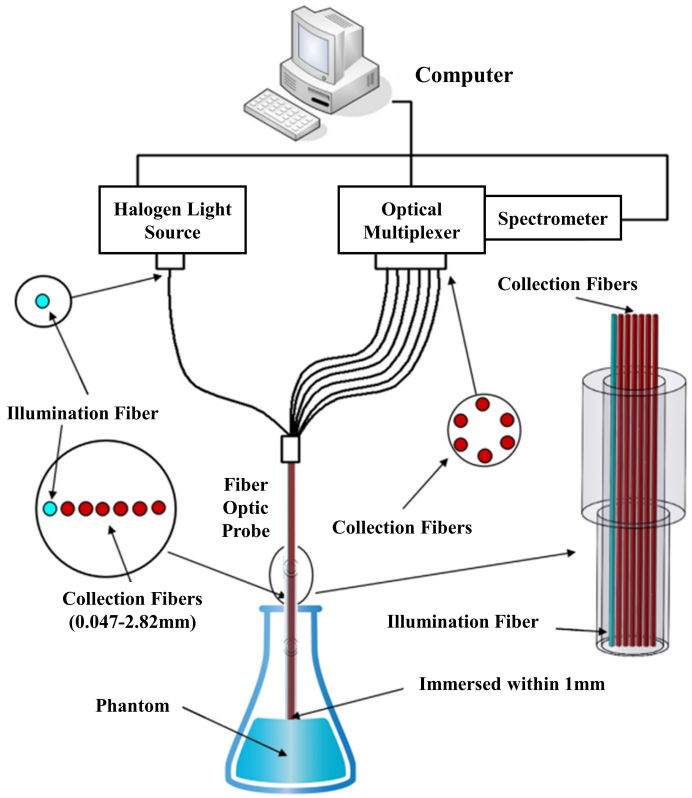


FIGURE 4
Fiber optic array type inspection device (Zhou et al., 2015).

camera, data acquisition, and control equipment. Spectral data were collected at intervals of 0.15 mm between the detection fiber and the illumination fiber over a range of 0.3-1.2 mm. The setup could split the diffuse light into multiple wavelengths in the range of 500-1000 nm by means of a spectrometer and project them onto different areas of a CCD camera (S7031-1008S, Hamamatsu, Japan). Finally, a data acquisition card and a customized LabView program (LabView 8.5, National Instruments, USA) were used to collect spatially resolved information.

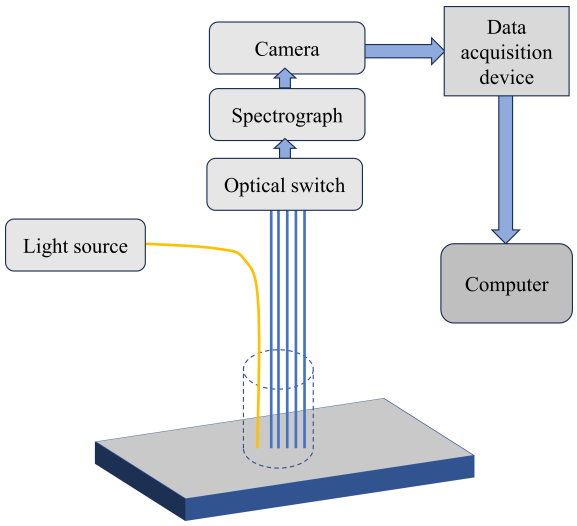


FIGURE 5
Novel fiber optic array device (Nguyen Do Trong et al., 2011).

In order to collect spectral information at more distances, Ma et al. designed a Vis-NIR SRS system (Ma et al., 2021b), as shown in Figure 6, which consists of a 5 W halogen light source, a Vis-NIR HSI camera and 30 silica fibers (core diameter: 100 μm, cladding: 110 μm), with five groups of fibers, each consisting of six fibers, 1, 2, 3, 4 and 5 mm away from the light source, respectively. The 30 silica fibers installed in this SRS acquisition device including both horizontal and vertical spatial-spectral information of the sample under test, which could increase the exploration of the spatially resolved spectral information.

These fiber optic array-type devices are mainly arranged in the form of linear arrays (Nichols et al., 1997; Doornbos et al., 1999; Bogomolov et al., 2017) and circular arrays (Dam et al., 2001; Nguyen Do Trong et al., 2011; Bridger et al., 2021). Their arrangement can be designed according to the sample's shape and structure's size. Since the designed fiber array structure is fixed to detect the spectral information at once, it can save the measurement time as well as avoid the spectral error caused by the inaccuracy of the distance during the measurement. However, custom-designed fibers are more costly and require testing and calibration of the fiber arrays. The fiber optic array is only suitable for detecting samples with a flat surface for most agricultural products due to the irregularity of the sample detection fiber probe not well fitted to the sample surface. In addition, this system requires contact with

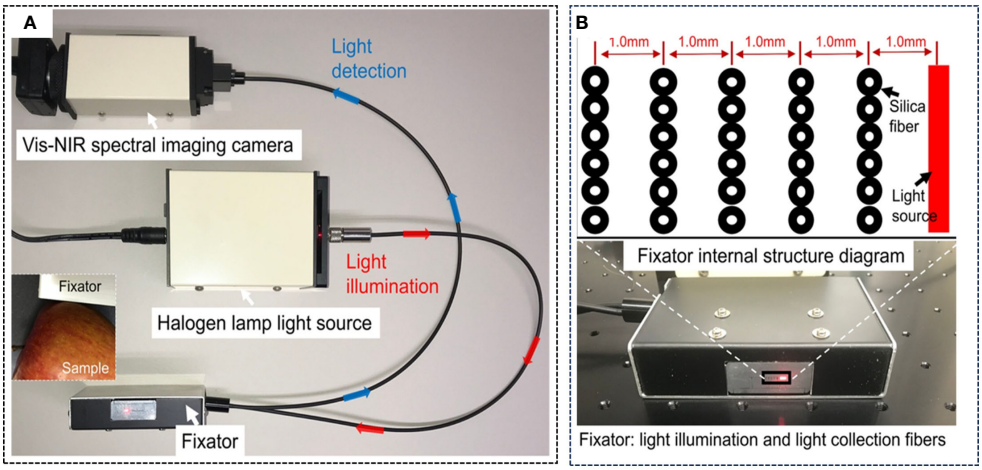


FIGURE 6
Vis-NIR SRS system (Ma et al., 2021b). **(A)** SRS detection systems. **(B)** Internal structure diagram of the fixator.

the sample surface during the inspection process and is not friendly to liquids, easily polluted and vulnerable samples, so it needs to be continuously improved and developed.

2.3 CCD line scan type system

The fiber optic array detection method is suitable for measuring liquid samples because the integrated array probe can make good contact with the liquid surface. In addition, it is also well suitable for flat sample pieces, such as dried apples or tablets (Igne et al., 2015), etc., but it is easy to contaminate the sample with this contact detection method, so it needs to be cleaned frequently. In order to achieve a non-contact measurement method while detecting the spatially resolved information of the sample, researchers have developed a spatially resolved detection system based on the CCD line scan method.

The spatially resolved system of CCD line-scan type is also called monochromatic imaging spatially resolved system, which is available for detecting the OP of a sample at a single wavelength. As shown in Figure 7, Kienle et al. (1996) used this approach for inspection. The system mainly consists of a laser diode as a light source, which is illuminated by a mirror at an angle of incidence of 5-10° on the object to be measured and detected by a CCD camera, and then the detected data are read out and processed by a computer.

Since laser diodes can only emit a single wavelength, this is not very friendly for analyzing multiple wavelengths. Therefore, researchers have pooled diodes at several different wavelengths for detection, which was used by Lorente et al. to detect the early

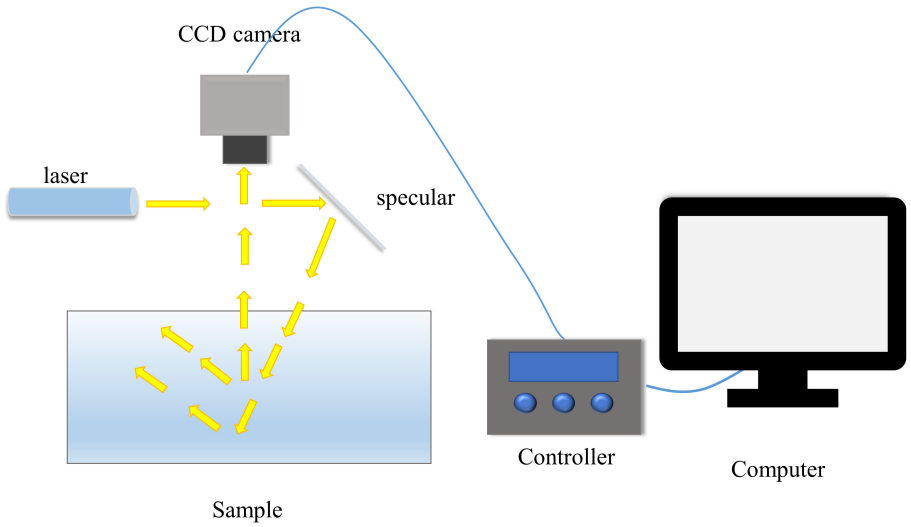


FIGURE 7
CCD line scan spatial resolution system (Kienle et al., 1996).

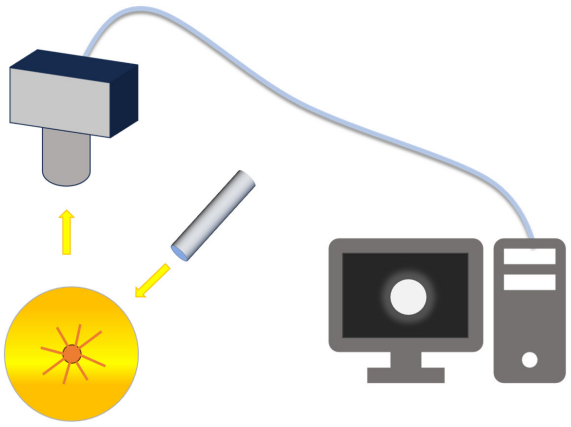


FIGURE 8
Laser diode-based optical properties device (Lorente et al., 2013).

ripeness of citrus fruits (Lorente et al., 2013). As shown in Figure 8, the system consists of a CCD camera, five solid-state laser diodes emitting at different wavelengths (532, 660, 785, 830, and 1060 nm), and a computer. In the acquisition process, the laser diodes are not integrated together for the acquisition, but the alternating form of replacing the diode of the corresponding wavelength each time to be used as a light source, so as to achieve the acquisition of spectral information at different wavelengths.

Conventional CCD imaging systems do not contain spectroscopic components inherently, and the acquired images are ordinary RGB images. Due to the theory of optics, only lasers or laser diodes can be used as light sources for CCD cameras. This limits the system to detecting the optical properties of the sample at a single wavelength. Although, at this stage, there is a way to detect spectral information in multiple wavelengths using diode module integration, it is still far from sufficient for analyzing continuous wavelengths. Moreover, the saturation of pixels occurs close to the light source point, so this area cannot be used for data analysis, and to avoid saturation, it is usually necessary to limit the exposure time (Kienle et al., 1996). This makes the CCD line-scan type system not well suited to the needs of the application, so a more optimized spatially resolved detection system is urgently needed.

2.4 Hyperspectral line-scan system

For SRS, the more continuous wavelength bands the collected information contains, the more advantageous it is likely to be for subsequent data analysis and processing. In pursuit of acquiring spatially resolved spectral information in continuous bands in a non-contact system, researchers have combined hyperspectral imaging (HSI) techniques with SRS, and they have been widely developed and applied. As shown in Figure 9, Peng and Lu (Peng and Lu, 2008) used a spatially resolved line-scan system, which mainly consists of a back-illuminated camera (C4880-21, Hamamatsu Photonics, Hamamatsu Corp., Japan), a control unit, an imaging spectrometer (ImSpector V9, Spectral Imaging Ltd., Oulu, Finland), a quartz tungsten halogen lamp (Oriel Instruments, Stratford, CT, USA) and a circular open sample holder with a diameter of 30 mm. The light source is a 1.5 mm circular beam, and the hyperspectral imaging system line scan is 1.6 mm from the light source to avoid oversaturation of the CCD detector pixels.

To make it easier to detect the OP of SRS, as shown in Figure 10. Cen and Lu (Cen et al., 2011) developed the Optical Property Analyzer (OPA), which consists of three main hardware components that are imaging, illumination, and sample positioning units. The imaging device mainly consists of an electron-multiplying CCD (EMCCD) camera (LucaEM R604, ANDORTM Technology, USA), an imaging spectrometer (ImSpector V10E, Spectral Imaging Ltd., Oulu, Finland), and a

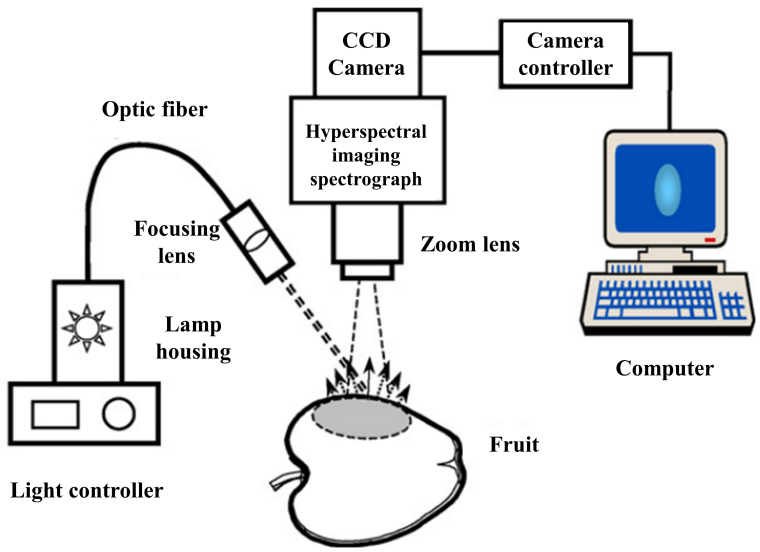


FIGURE 9
Hyperspectral imaging system for acquiring spectral scattering images (Peng and Lu, 2008).

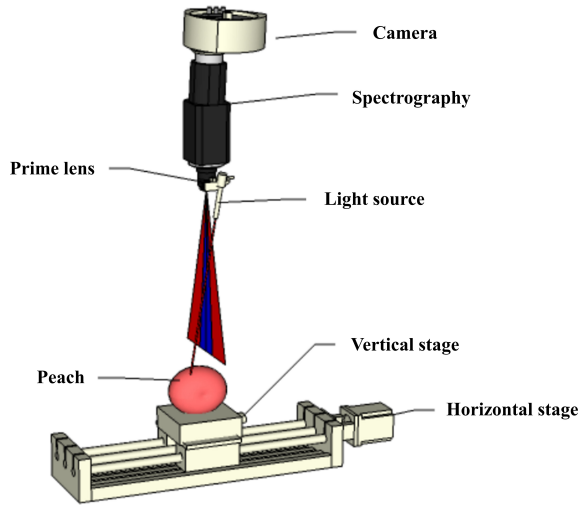


FIGURE 10
The Optical Property Analyzer (OPA) (Cen et al., 2011).

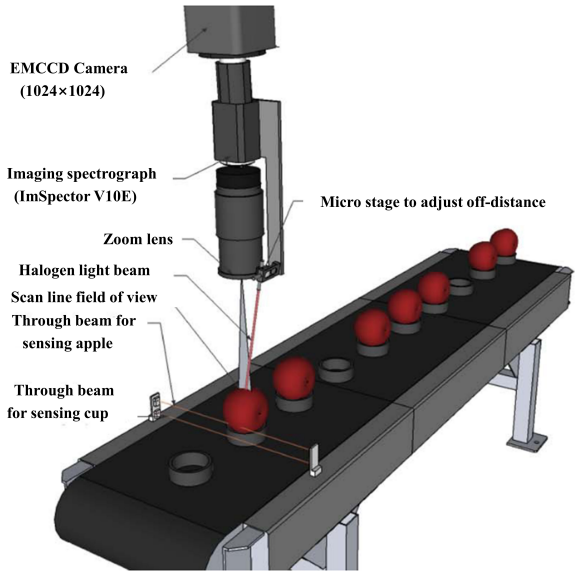


FIGURE 11
Online hyperspectral imaging system (OHIS) (Mendoza et al., 2011).

master lens (Xenoplan 1.9/35, Schneider Optics, Hauppauge, USA). An optical fiber connected to the focusing lens can be used to emit a point light source. The sample fixation device consists of a motorized horizontal stage (Twintrac, TSZ8020, US23T22104-8LS, US Automation, USA) with a maximum speed of 203 mm/s and a positioning accuracy of 0.0006 mm/mm, a vertically adjustable stage, and a holder for sample positioning. The integrated software program for OPA is developed in Microsoft Visual C#. It can control the light source, camera, and sample mounting platform for spectral and image data acquisition and also analyze and display the acquired information in real time to obtain the final scattering profile, absorption spectrum, reduced scattering coefficient, etc. Due to the powerful and convenient functions of this software, the workload of spectral data acquisition and analysis can be greatly reduced, and the efficiency of the sample acquisition and analysis can be improved.

Mendoza et al. (2011) developed an online hyperspectral imaging system (OHIS) based on a hyperspectral line-scan type (Figure 11), which consists of a back-illuminated EMCCD camera, an imaging spectrometer (ImSpector V10E, Spectral Imaging Ltd., Oulu, Finland) covering a spectral region of 400-1000 nm. A near-infrared enhancement lens and a halogen light source (Oriel Instruments, USA). The computer is equipped with an image acquisition card and a camera acquisition program written in C++, through which the camera can be controlled for image acquisition. In order to capture the samples in real time and to increase the efficiency of the test, the device also uses a conveyor belt that can hold the samples. The imaging system of this OHIS operates at a rate of approximately one in two seconds. This system is the first to combine spatially resolved line sweep with online inspection. Although the system has good predictions, it is costly and still has some errors for curved samples.

While hyperspectral imaging inspection methods can realize the advantages of contactless, efficient, and high-resolution acquisition, they also have significant drawbacks. However, it is only suitable for detecting samples with flat surfaces or objects of considerable size, and if the surface curvature of the sample is large, the detected information will have a large error, so the detection device of SRS needs to be improved continuously.

2.5 Multi-channel hyperspectral imaging detection system

The current device is only suitable for detecting samples with relatively flat surfaces, and the detection probes cannot fit closely for most agricultural products. When the detection sample is too large, the existing fiber array system makes it difficult to meet the requirements of the number of detection fibers and detection distance due to the limitations of the instrument. Although hyperspectral detection has excellent advantages, it has a narrow detection wavelength range and lacks flexibility for curved samples, which can cause significant errors. Therefore, Huang et al. (2017) designed a multichannel hyperspectral imaging detection device, as shown in Figure 12, which was based on a multichannel hyperspectral imager (Headwall Photonics, Inc., USA). The multichannel probe consists of a point source and 30 fibers of three sizes (i.e., 50 μm, 105 μm, and 200 μm). The light source fiber is connected to a 250 W halogen lamp, and the 30 fibers are permanently mounted on two sizes of aluminum cubes, giving the probe the flexibility to measure samples of different sizes and flat or curved surfaces at distances of 1.5-36 mm.

In general, the proposed SRS detection devices have their own advantages and disadvantages as well as applicable detection samples. The characteristics of these detectors are described in Table 1. Although the single fiber detection system is simple in structure, lowest cost, and flexible in collection, there will be a large measurement error and time-consuming, so it is not suitable for a large number of sample collections. Fiber array type system can

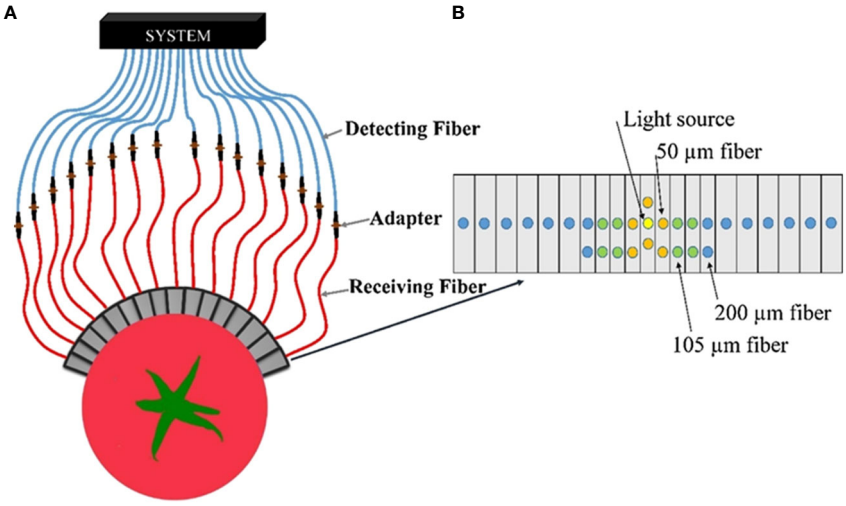


FIGURE 12
Multichannel hyperspectral imaging detection device (Huang et al., 2017). **(A)** Schematic of spatially resolved spectral acquisition. **(B)** Schematic of fiber arrangement.

TABLE 1 Summary of studies on the types of detection devices for agricultural products.

Detection systems	Objects	Wavelength(nm)	Detection distance (mm)	Characteristics	References
SF	Beef	490-950	Incident fiber Left: 9.0-6.5 Right: 4.0-7.0	Detection flexibility allows the selection of the optimal distance	(Xia et al., 2007)
	Apple	190-1070	2, 4, 6, 8, 10, 12, 14, 16	Easy operation, reduce error	(Ye et al., 2021)
	Pear	500-1000	-0.15, -0.1, -0.05, 0, 0.05, 0.1, 0.15	–	(Hu et al., 2017)
	Onion	710-950	–	The laser system has a slightly better optimal single point ratio than the NIRS system	(Sun J. et al., 2020)
	Rabbit	350-1000	5, 10, 15	Detecting distance slidable	(Yuan et al., 2022)
FA	Milk	550-1690	1-2.5	Optimal combination of minimum fiber counts	(Watte et al., 2016)
	Milk	400-995	0.28-1.96	Full-spectrum analysis replaced by two wavelength-specific sensor measurements	(Bogomolov et al., 2017)
	Pork	600-1100	6, 9, 12, 15	Improve detection efficiency	(Wen et al., 2010; Zhang et al., 2010)
	Pork	600-1100	6, 9, 12, 15	Efficient, low cost	(Wang J. et al., 2017)
	Apple	500-1000	0.3-1.2	Efficient	(Nguyen Do Trong et al., 2014a; Nguyen Do Trong et al., 2014b)
	Apple	600-1100	1, 2, 3, 4, 5	Portable, high efficiency	(Ma et al., 2021b)
	Kiwifruit	660-1000	1, 2, 3, 4, 5	Portable, high efficiency	(Ma et al., 2022)
	Wood	600-1100	2, 3, 4, 5	Portable, high efficiency	(Ma et al., 2021c)
	Cattle	500-900	0.5, 1.0, 1.5, 2.0, 2.5	Fiber integration, high efficiency	(Palendeng et al., 2020)

(Continued)

TABLE 1 Continued

Detection systems	Objects	Wavelength(nm)	Detection distance (mm)	Characteristics	References
CL	Milk	800-1065	–	Fast, portable and inexpensive	(Kalinin et al., 2013)
	Apple	650-980	–	Detecting distance slidable	(Mollazade and Arefi, 2017)
	Banana	532, 660, 785, 830, 1060	–	Specific wavelength, non-contact	(Adebayo et al., 2016)
	Citrus	532, 660, 785, 830, 1060	–	Specific wavelength, non-contact	(Lorente et al., 2013)
	Wood	808	Dry: Parallel: 20 Perpendicular: 10 Wet: Parallel: 25 Perpendicular: 15	Non-contact	(Kienle et al., 2008)
HL	Milk	530-900	1.6-20	Non-contact, easy operation	(Qin and Lu, 2007)
	Apple	500-1000	1.6-9	Easy to operate	(Qin et al., 2007; Qin et al., 2009; Lu et al., 2010)
	Apple	450-1000	–	Easy to operate	(Peng and Lu, 2008)
	Apple	600-1000	–	Easy to operate	(Huang and Lu, 2010)
	Apple	500-1000	0-9	Easy to operate, with analysis software	(Cen et al., 2012b; Cen et al., 2013)
	Apple	450-1050 (Scatter) 460-1100 (Vis/SWNIR)	–	Realized hyperspectral online detection	(Mendoza et al., 2014)
	Apple	500-1000	20	Easy to operate, with analysis software	(Zhu et al., 2016)
	Peach	550-1650	1-9	Easy to operate, with analysis software	(Cen et al., 2011; Cen et al., 2012a)
	Peach	550-1000	1-9	Easy to operate, with analysis software	(Sun Y. et al., 2020; Sun Y. et al., 2021)
	Cucumber	700-1000	37-55	Easy to operate, with analysis software	(Lu et al., 2011)
	Tomato	500-950	0-10	Easy to operate, with analysis software	(Zhu et al., 2015)
	Wood	1000-1600	1, 3, 5 (Thicknesses)	Non-contact methods, push-broom manner	(Ma et al., 2018; Ma et al., 2019b)
	Wood	1002-2130	–	Non-contact methods, push-broom manner	(Ma et al., 2019a)
	Tea	967.11-1700	–	Non-contact methods, push-broom manner	(Mishra et al., 2019)
MHI	Apple	550-1650	1.5-36	Detects longer distances and higher accuracy	(Huang et al., 2020b)
	Peach	550-1650	1.5-36	Detects longer distances and higher accuracy	(Huang et al., 2022)
	Tomato	550-1300	1.5-36	Detects longer distances and higher accuracy	(Huang et al., 2018b) (Huang and Chen, 2018)
	Tomato	550-1650	1.5-36	Detects longer distances and higher accuracy	(Huang et al., 2018c; Huang et al., 2020a)

SF, Single fiber system; FA, Fiber array type system; CL, CCD line-scan type system; HL, Hyperspectral line-scan system; MHI, Multi-channel hyperspectral imaging detection system.

achieve once-time acquisition at different distances to improve detection efficiency and accuracy, but the cost is higher than single-fiber detection systems with the need to detect samples as smoothly as possible, and contact measurements are prone to sample contamination, so the scope of use is also very limited. CCD line scan type system can realize non-contact measurements but cannot collect SRS in the continuous wavelength band. The hyperspectral line-scan system can acquire spectral information in continuous wavelength bands and are well suited for the detection of a large number of samples, while they are less friendly to the detection of samples with curvature, such as apples, peaches, oranges, etc. Multi-channel hyperspectral imaging detection system can detect the spectra of some curvature samples, but the cost is the highest, and there is still a large error for some samples with large curvature. In general, although these systems have detected the spatially resolved information of samples to a great extent, they are still not friendly enough for the detection of irregular objects because of their detection limitations, so there is

still a huge space that could be developed and innovated for the detection systems of SRS.

3 Development of spatially resolved spectral analysis methods

SRS collects spectral information at different distances and then needs to be processed. In most cases, the optical properties are obtained based on SRS techniques, and then the association between sample properties and optical properties is analyzed on the basis of absorption coefficients and reduced scattering coefficients. Direct analysis is also used to detect the properties by processing the spectral information at different distances. At this stage, researchers have done a lot of work on the basis of optical properties and direct analysis, which provides powerful and effective support for the development of SRS in the future.

3.1 Optical properties methods

Among the methods of measuring OP in biological tissues, there are currently two ways: direct and indirect measurement, respectively. In the direct measurement method, Beer-Lambert's Law is used to calculate the OP within the tissue. In this method, the optical properties of the tissue are calculated by measuring parameters, such as complete attenuation transmission and collimated transmission of a slice sample. Although the direct analysis calculation method is simpler, its detection process is more complex, requiring slicing and strict requirements for the thickness of the slice (Cheong et al., 1990). The indirect measurement method is mainly used to solve the OP by inversion. Generally, researchers classify indirect measurements into non-iterative and iterative approaches according to whether the inversion process includes a parameter iteration step. The non-iterative approach can be used to solve the optical properties of the optical transmission model directly from the measured values. One of the more commonly used methods is the Kubelka-Munk method (Kubelka, 1948), but the accuracy of its measurements is not high, requiring assumptions on various conditions. The iterative approach is to evaluate the OP by inverting the parametric equations for several iterations so that the measured values are within the specified error range. Although this method is more complicated, the measured optical properties are more accurate than other methods. Spatially resolved techniques are also usually applied by using indirect iterations to find the OP within the sample tissue.

For the transmission of light in biological tissues, a series of complex optical phenomena occurred, such as absorption, scattering, reflection, refraction, interference, and diffraction of light. Although Maxwell's set of equations based on electromagnetic theory can describe the light propagation process in tissues mathematically (Yang et al., 2021; Katsumata, 2022), the equations cannot be solved directly due to the complexity of biological tissues. In order to study only the particle properties, such as absorption and scattering of light, the fluctuating properties of light, such as interference, diffraction, and polarization, can be ignored. Researchers have proposed the Radiative Transfer Equation (RTE), which is more accurate in describing the transport properties of light in tissues (Martelli et al., 2021; Hank et al., 2023), but the model is still complex and has many parameters. Therefore, researchers usually have used simplified transport model methods as well as numerical methods to solve the optical properties (Lu et al., 2020). The commonly used transport models are the diffusion approximation and the P3 approximation, which is the theoretical model to describe the spatially resolved diffuse reflection near the light source. The P3 approximation model is a third-order form of the radiative transfer model. Since the P3 approximation is more accurate, it can be used in place of the diffusion approximation (Wang, 2020; Wang, 2022). In addition, numerical methods include Monte Carlo (MC) simulation (Chong and Pramanik, 2023; Colas et al., 2023; Lan et al., 2023), Adding-Doubling model (Xie and Guo, 2020; Sun et al., 2022) and finite element methods (Morimoto et al., 2020).

The diffuse approximation equation is a simplified form of the radiative transfer equation, which has the ability to be simplified by satisfying two assumptions. 1) The medium is a strongly scattering medium, i.e. $\mu_s{}' \gg \mu_a$. 2) The SDD is greater than the mfp' (mean free path), i.e. r>mfp'. In addition, the incident light scattering step in the tissue is considered to be isotropic radiation. The diffuse approximation equation can be expressed as:

$$\frac{\partial \boldsymbol{\Phi}(\vec{r},t)}{c\,\partial t} + \boldsymbol{\mu_a}\boldsymbol{\Phi}(\vec{r},t) - \nabla\cdot[\boldsymbol{D}\nabla\boldsymbol{\Phi}(\vec{r},t)] = \boldsymbol{S}(\vec{r},t)$$

where c is the spreading speed of light through the medium, is the radiation fluence rate, $\vec{r} = (x, y, z)$ is a point within the medium, $D = \left[3\left(\mu_a + \mu_s^{'}\right)\right]^{-1}$ is the diffusion coefficient, $S(\vec{r}, t)$ is each homogeneous light source. This equation can be used to describe the transmission of light through some objects with geometric shapes, such as semi-infinite, flat, cylindrical, spherical, etc. (Farrell et al., 1992; Kienle et al., 1998), which provides a good application for detection of OP of most samples. Depending on different illumination methods such as steady-state point illumination, pulsed point illumination, frequency-modulated point illumination, and spatially modulated area illumination, OP techniques have also evolved into spatially resolved techniques, time-resolved techniques, frequency domain resolved techniques, and spatial frequency domain techniques (Lu et al., 2020).

Based on the theory of diffuse approximation equations, Farrell et al. proposed a diffusion-theoretic model for SR steady-state diffuse reflection in the study of nondestructive determination of OP in humans (Farrell et al., 1992). The model can be used to describe the directional dependence of light diffuse reflection in biological tissues when irradiated by an infinitesimal amount of light. By comparing the predictions of the model with MC simulations and with tissue simulation models, it was found that the model can accurately describe the reflectance at radial distances as small as 0.5 mm. Thus, the model can provide an effective method and basis for later researchers to calculate and separate the OP. In this model, the diffuse reflection of the medium is computed as a boundary flow from a single isotropic point source located at the mfp' depth of the medium transport. The model is applied to surfaces with matched or unmatched refractive indices. The

equations of this diffusion model are as follows:

$$R(r) = \frac{a'}{4\pi}\left[\begin{array}{l}\frac{1}{\mu_t}\left(\mu_{eff}+\frac{1}{r_1}\right)\frac{exp\left(-\mu_{eff}r_1\right)}{r_1^2}+\\ \left(\frac{1}{\mu_t}+\frac{4A}{3\mu_t'}\right)\left(\mu_{eff}+\frac{1}{r_2}\right)\frac{exp\left(-\mu_{eff}r_2\right)}{r_2^2}\end{array}\right]$$

where r is the source-detector distance, $r_1 = \left(z_0^2 + r\right)^{1/2}$ is the real distance from the detector to the light source, $r_2 = \left((z_0 + 2z_b)^2 + r^2\right)^{1/2}$ is the length from the mirror light source to the detector. $\mu_{eff}' = \left[3\mu_a\left(u_a + \mu_s'\right)\right]^{1/2}$ is the effective reduction factor, $a' = \mu_s'/\left(\mu_a + \mu_s'\right)$ is an albedo of transmission, $\mu_t' = \mu_a + \mu_s'$, is the overall decays value, $z_0 = \left(\mu_a + \mu_s'\right)^{-1}$ is the mfp', $z_b = 2AD$, A is the object's internal reflection coefficient, A=1 when the tissue and surrounding media boundaries match, and A=0.2190 when the relative refractive indices of the tissues n=1.35. While refractive index is known to be wavelength dependent, most reports assume that the n is constant, an assumption that is subject to potential uncertainty, such as for many fruits and foods n=1.35.

Later, Kienle and Patterson (1997) introduced radiant energy flow rate following Haskell et al. (1994). The diffuse reflectance is expressed through the brilliant energy flow rate and luminous flux, which better minimizes errors and thus more precisely characterizes the transmission of light in biological tissues. This equation can be expressed as:

$$\Phi(r, z = 0) = \frac{1}{4\pi D}\left[\frac{exp\left(-\mu_{eff}r_1\right)}{r_1} - \frac{exp\left(-\mu_{eff}r_2\right)}{r_2}\right]$$

The diffusive approximation equation can be organized as follows:

$$R(r) = \frac{C_1}{4\pi D}\left[\frac{exp\left(-\mu_{eff}r_1\right)}{r_1} - \frac{exp\left(-\mu_{eff}r_2\right)}{r_2}\right] +$$

$$\frac{C_2}{4\pi}\left[\frac{1}{\mu_t'}\left(\mu_{eff}+\frac{1}{r_1}\right)\frac{exp\left(-\mu_{eff}r_1\right)}{r_1^2} + \left(\frac{1}{\mu_t'}+2z_b\right)\left(\mu_{eff}+\frac{1}{r_2}\right)\frac{exp\left(-\mu_{eff}r_2\right)}{r_2^2}\right]$$

where $C_1 = \frac{1}{4\pi}\int_{2\pi}\left[1 - R_{fres}(\theta)\cos\theta d\omega\right]$ and $C_2 = \frac{3}{4\pi}\int_{2\pi}\left[1 - R_{fres}(\theta)\cos^2\theta d\omega\right]$ are coefficients generated by the refractive index of the medium and $R_{fres}(\theta)$ is the Fresnel coefficient. When rate of refraction n=1.35, C1 and C2 are 0.1277 and 0.3269, respectively (Cen and Lu, 2010). This solution is considered to be more accurate in describing the light propagation process, therefore it is widely used (Cen et al., 2010).

At present, these two are the most commonly used mathematical fitting models for spatially resolved spectra, and their analytical solutions are obtained under extrapolation boundary conditions (EBC). The source is assumed to be each homogeneous radiation source at one mfp' below the sample surface. Therefore, the precision of OP parameter inversion is not only related to the precision of instrumental measurements but also depends on the precision of parameter inversion algorithms. Cen and Lu further optimized the curve fitting algorithm by using the nonlinear least squares method as the Trust-region-reflective least squares method, and the raw data were logarithmically and integrally transformed and relatively weighted before fitting to improve the OP predictions (Cen et al., 2010). The prediction of the OP was enhanced by using logarithmic and integral transformations of the original data and relative weighting before fitting.

However, many factors affect the accuracy and error in the acquisition and processing of the spectra and the inversion of the parameter equations. Usually, normalization is required before curve fitting. However, the standard normalization method directly divides the first value of the spatially resolved diffuse reflectance spectrum, which contains considerable noise and acquisition errors. It has a great impact on the inversion of the later parametric equations. The diffuse approximation equation is invalid when it is close to the light source, i.e. (r<1 mfp'), and the reflection signal is weaker, and the signal-to-noise ratio (SNR) is lower when the acquisition is farther away, which is not conducive to the inversion of the optical parameters. Therefore, an effective interval selection for the acquired spatially resolved spectra is also needed before curve fitting. Farrell et al. suggested that the SDD should be greater than one mfp' (Farrell et al., 1992), and Nichols et al. recommended minimum and maximum distances of SDD are 0.75-1 mfp' and 10-20 mfp', respectively (Nichols et al., 1997). Nevertheless, for most of the unknown samples with unknown OP, it is impossible to calculate the mfp' directly. So Wang and Lu et al. proposed a mean normalization method to optimize the normalization along with a method to optimize the diffuse reflectance spectral interval for the inversion of OP based on the relative error contour (Wang A. et al., 2017).

The inversion of the parametric equations is performed by fitting the diffuse reflectance spectral data to an analytical solution of the diffuse reflectance approximation equations to calculate the absorption coefficients and the reduced scattering coefficients. Cen and Lu used the spectral SNR to optimize the endpoint of the spectrum (Cen and Lu, 2010), but the starting point of the spectrum is fixed at 1.5 mm for systematic reasons, which still leads to large measurement errors for samples with mfp' greater than 1.5 mm measurement error. Therefore, to further solve the problems of fitting, Wang and Lu proposed the step-by-step parameter inversion method (Wang A. et al., 2017), which is based on the OP and the mfp' obtained by the one-step fitting method, and then re-determine the better spectral interval based on the optimized starting point and end point before the second fitting to obtain the better OP. The method is effective in improving the optical parameters. This method can significantly improve the inversion accuracy of optical parameters.

However, in the process of inversion, the traditional inversion algorithm does not meet the requirements due to the single-layer and double-layer tissues of the sample under test. The traditional inversion algorithm is to equate the outer tissue of the sample with the inner tissue as a layer. For samples with a thin outer skin, the effect of having a thickness less than the mfp' will not be significant, but for pieces with a thicker outer skin, the effect will be more meaningful if they are equated to a monolayer of tissue. There are already diffuse reflection equations for light transmission in single-

and double-layer tissue that can be used as an approximate solution (Kienle et al., 1998; Cen and Lu, 2009), but the light transmission in double-layer tissue is a more complex transmission model involves more parameters, the accuracy of the fitted parameters will be increasingly poor. Accordingly, the inversion of parametric equations for multilayer tissues still needs to be continuously explored and studied by researchers.

The most commonly used method in numerical calculation methods is MC, which is a statistical method with random sampling and has been widely used to simulate the propagation of light (Tarasov et al., 2021; Sassaroli et al., 2022). This method can simulate the light transmission process by tracking the trajectory of a massive photon through the tissue and finally calculate the optical parameters we need. The advantages of MC are low cost, high accuracy, and high flexibility. However, it also has obvious drawbacks, which are computationally intensive, time-consuming, and not conducive to rapid detection, so the method is often used to test the accuracy of other calculation methods. When the MC method simulates the light transmission law in tissues, it mainly simulates the particle properties of light, i.e., the absorption and scattering of light and other properties. Its typical simulation specifically includes the processes of photon generation, initialization, migration, absorption, scattering, boundary condition processing, and extinction judgment (Wang et al., 1995). Currently, the program developed by Wang and Jacques (1992) based on C language can be used for multilayer organization, which consists of two subroutines, Monte Carlo Multi-Layered (MCML) and Convolution (CONV), where MCML is used to simulate the transmission of light beams in the organization. CONV is used to convolve the simulated data of MCML and output the results. Based on the disadvantages of MC time consumption, Hu et al. (2020b) optimized it and accelerated its simulation. Sun et al. used the Monte Carlo multilayer (MCML) technique to simulate the propagation of light through the fruit by comparing it with the diffuse reflection curve, thus confirming the accuracy of the MC simulation of the OP (Sun C. et al., 2021).

Although MC is usually used as a reference method and is more accurate, it needs a huge number of photons to be simulated at a time, which is computationally intensive and cannot meet the rapid detection of OP of biological tissues despite the fact that its speed has been improved. The finite element method (FEM) is also one of the commonly used numerical methods, which is more flexible and fast based on accuracy (Vasudevan and Narayanan Unni, 2021). Lee et al. (2004) used the FEM method to study the propagation of light in a double-layer medium and found that the accuracy of the method and MC were almost the same by comparison. Wang et al. (2016) investigated the optimal computational results of finite elements under three boundary conditions and demonstrated that the finite element approach can be used to improve the measurement of OP for spatially resolved techniques. Whereas, at this stage, there are few analytical methods using finite elements in the study of calculating the OP of spatially resolved spectra, which is a promising method for the numerical calculation of OP.

3.2 Direct analysis methods

While analysis methods based on OP can find absorption coefficients and approximate scattering coefficients more accurately, they are labor-intensive, algorithmically complex, and have large accuracy errors. In order to directly and accurately analyze spatially resolved (SR) spectra as well as to simplify the analysis steps, the researchers proposed a direct analysis method.

Huang and Chen (2018) proposed an analytical method of spectral combination when employing a multichannel SRS system to detect tomatoes by creating a Partial Least Squares Discriminant Analysis (PLSDA) model of each of the 15 single SR spectra combinations to determine the best single SR combination for classification. Next, the best SR combination was combined with the remaining 14 SR combinations to select the best two-spectrum combination, then the best spectrum was combined with the remaining 13 single SR combinations to create the best tri-spectral combination, and so on until the accuracy of the combined SRS for classification is not further improved.

As for quality detection of peach, Huang et al. (2022) proposed a spectral difference technique to deal with spatially resolved spectral information. The method initially collects 30 relative spatial spectra for each sample at different SDD scales, which are calculated as follows:

$$\boldsymbol{R(i)} = \frac{\boldsymbol{I_s(i)} - \boldsymbol{D_s(i)}}{\boldsymbol{I_r(i)} - \boldsymbol{D_r(i)}}$$

where R is the relative spectrum, I is the spectral information of the sample, D is the blackboard, i is for single fiber, $i = 1, 2, 3,\ldots, 30$, and the subscripts r and s represent the white Teflon and the sample, respectively. Since the device has detection fibers arranged symmetrically, each symmetrical pair of SR spectra is averaged over the same SDD, resulting in 15 SR spectra whose distances range from 1.5-36 mm. The difference spectrum is obtained by differencing the spatially resolved spectrum of the first position (SR1) with the spatially resolved spectra of the other SDDs with the following calculation equation:

$$\boldsymbol{D(i)} = \boldsymbol{R(i)} - \boldsymbol{R(1)}, \ \ \boldsymbol{i = 2, 3, 4, 5, \cdots, 15}$$

where $R(1)$ is the SR spectrum at the first SDD of 1.5 mm, and D is the subtraction spectrum, after that, it is referred to as the differential reflectance (DR) spectrum, and the final 14 DR spectra contain different spatial resolution information compared to the SRS.

Ma et al. (2021b) proposed a method for reference-free reflectance calculation in assessing apple quality by averaging the light centers at distances of d mm and $d+\Delta$ mm for the diffuse reflected light intensity (i) collected by the optical fibers and named i_d and $i_{d+\Delta}$, respectively. When the intensity of the spatially resolved spectrum is i_{ref}, the spectral difference between fibers at different distances can be expressed as:

$$\boldsymbol{A_{diff}} = -\boldsymbol{log}_{10}\left(\frac{\boldsymbol{i_{d+\Delta}}}{\boldsymbol{i_{ref}}}\right) - \left(-\boldsymbol{log}_{10}\left(\frac{\boldsymbol{i_d}}{\boldsymbol{i_{ref}}}\right)\right)$$

where A_{diff} is the difference in absorption spectra. The equation can be simplified as:

$$A_{diff} = -log_{10}\left(\frac{i_{d+\Delta}}{i_d}\right)$$

In this way, the calculation formula of the spectrum can be simplified, and the black-and-white correction of the spectrum can be canceled, which makes the spectral inspection more efficient and convenient. The ratio of diffuse light intensity (R_{Ratio}) is calculated as follows:

$$R_{\mathrm{Ratio}} = \frac{i_{d+\Delta}}{i_d}$$

Finally, smoothing of the spectral data using Savitzky-Golay filters enables the spectra to achieve better results in modeling analysis.

The direct analysis method of SRS simplifies the analysis steps, and although it is not more accurate than the OP method, it has the same good prediction effect for the quality detection of agricultural products. There are few direct analysis methods used so far. If a better direct analysis method can be proposed to predict the quality of products, not only the analysis method is simple and fast, but also the quality prediction accuracy is more accurate, then the detection efficiency of SRS will be significantly improved.

4 Application of spatially resolved spectroscopy in agricultural products

Although SRS has been widely adopted in the biomedical field, its application in the agricultural field is still relatively limited. At present, the application of spatially resolved technology is mainly concentrated in the field of edible agricultural products, such as meat, dairy, fruits, and vegetables, and less application in other areas, such as forestry, animal husbandry, etc. The technology of detecting the quality or classification of agricultural products by SRS is more mature. In the subsequent sections, the latest research and specific applications of SRS in agriculture were presented and summarized in detail.

4.1 Applications of dairy field

In the field of dairy products, SRS is more widely used in the detection of milk. Because milk is rich in nutrients such as protein, fat, vitamins, and minerals, it is very easy to be absorbed by the human body, so it is very popular among human beings. To ensure the quality of raw milk or to prevent adulteration during the sale process, it is necessary to test the quality of raw milk. Watte et al. developed a global optimizer that can calculate the optimal configuration of fibers, by which the number of detected fibers can be minimized while maintaining the validity of the OP evaluation, making the detection optimal with cost savings. The design achieved good results for the evaluation of the OP of milk, with a root mean square error of the prediction (RMSEP) of 0.382 cm^{-1} and R^2 = 0.996 for the reduced scattering coefficient values (Watte et al., 2016). Kalinin et al. used a dual-channel short-wave near-infrared spectrometer as a detection device. The results showed that the RMSEP of proteins using a combination of scattering and transmission spectroscopy could reach 0.25% wt. (Kalinin et al., 2013). Bogomolov et al. developed and utilized a fiber-optic array-based detection device with eight channels of probes to analyze the quality of milk and improve the accuracy of fat and protein detection, The root mean square errors (RMSE) for the different validation methods were less than 0.10% for fat content and less than 0.08% for total protein content, respectively. (Bogomolov et al., 2017). The optimal sensor configuration was proposed to replace the full spectrum analysis with LED in specific wavelength bands, which provided a faster and more mature application for milk detection. Qin and Lu used a hyperspectral line-scan detection device to analyze the fat content in milk. They found that the absorption coefficient and the reduced scattering coefficient at 600 nm were closely correlated with the fat content of milk, while the R^2 were 0.995 and 0.998, respectively, which verified the feasibility of HSI in detecting the milk content (Qin and Lu, 2007).

As shown in Tables 1 and 2 in dairy product testing, researchers have used different spatially resolved detection systems to detect milk's fat and protein content to achieve good prediction results. However, more milk is currently detected, and the approach will definitely be developed toward a broader range of dairy products in future applications.

4.2 Applications of meat products field

In the detection of meat products, Xia et al. applied the SRS technique to the detection of meat products for the first time. They measured the SRS of beef samples with a single-fiber detection, obtained the absorption coefficient and scattering coefficient of beef through the diffuse reflectance equation, and established a correlation analysis between beef shear force and scattering coefficient, with a coefficient of determination (R^2) of 0.59, which verified the feasibility of SRS in detecting beef tenderness (Xia et al., 2007; Xia et al., 2008). Zhang et al. studied the tenderness of pork using multi-channel SRS and predicted the tenderness of pork by decreasing the scattering coefficient, which was R^2 = 0.8349 for fresh meat shear, through which the tenderness of pork can be directly predicted to realize fast and non-destructive detection (Zhang et al., 2010). Wen et al. investigated myoglobin content in pork and found that SRS in the short wave range was a feasible method for detecting myoglobin content with a significant correlation R^2 = 0.955 (Wen et al., 2010). Wang et al. determined the moisture content of the complete pork using SRS and found that the steady-state SRS was capable of significantly forecasting the moisture content of the pork compared to the conventional Y-fiber, with an R^2 of 0.8078 for their model. (Wang J. et al., 2017).

In summary, SRS is currently applied to detect tenderness, myoglobin and moisture content of meat products. Moreover, this technique can improve the accuracy of meat quality prediction to a great extent. Table 2 summarizes in detail the results of research on meat product quality testing. While relatively few meat products

TABLE 2 Summary of studies about the quality detection of agricultural products.

Products	Species	Applications	Methods	Accuracy	References
Dairy	Milk	Fat and protein contents	PLS	$RMSEP_f \leq 0.08\%$ $RMSEP_p$=0.21%	(Kalinin et al., 2013)
	Milk	Fat and protein contents	GA	μ_a:R^2 = 0.965 μ_s':R^2 = 0.996	(Watte et al., 2016)
	Milk	Fat and protein contents	PLSR, JVSPO	$RMSEP_f$<0.10% $RMSEP_p$<0.08%	(Bogomolov et al., 2017)
	Milk	Fat content	PLS	μ_a: R^2 = 0.995 μ_s': R^2 = 0.998	(Qin and Lu, 2007)
Meat	Beef	Tenderness	–	p<0.0001, R^2 = 0.59	(Xia et al., 2007)
	Pork	Tenderness	–	μ_s': R^2 = 0.8349	(Zhang et al., 2010)
	Pork	myoglobin	–	R^2 = 0.955	(Wen et al., 2010)
	Pork	moisture content	SPA, PLSR	R^2 = 0.8078	(Wang J. et al., 2017)
Fruit	Apple	Firmness and SSC	MLR, LCV	F: r=0.88, SEP=5.66N SSC: r=0.82, SEP=0.75%	(Qin et al., 2007)
	Apple	Firmness and SSC	MLR, MLD	Firmness: R=0.894, SEP=6.14 N; SSC: R=0.883, SEP=0.73%	(Peng and Lu, 2008)
	Apple	Firmness and SSC	MLR, LCV	Firmness: R=0.844 SSC: R=0.864	(Qin et al., 2009)
	Apple	Bruise detection	–	–	(Lu et al., 2010)
	Apple	Mealiness	PLS-DA	Accuracy>93%	(Huang and Lu, 2010)
	Apple	Firmness and SSC	PLSR	F: r_{GD}=0.892, r_{RD}=0.863 SSC: r_{GD}=0.892, r_{RD}=0.863	(Cen et al., 2012b)
	Apple	Mechanical and structural properties	ANOVA, LSD	Acoustic/impact firmness GD: r=0.870–0.948 GS: r=0.334–0.993 Young's modulus GD:r=0.585–0.947 GS: r=0.292–0.694	(Cen et al., 2013)
	Apple	Quality grades: firmness, SSC	LDA	Scattering technique Firmness: 77.9%-98.2% SSC: 62.0%-91.7% Vis/SWNIR technique Firmness: 87.3-97.6% SSC: 77.10-92.3%	(Mendoza et al., 2014)
	Apple	Microstructure, textural quality	–	–	(Nguyen Do Trong et al., 2014b)
	Apple	Firmness and SSC	PLS	Firmness: R^2 = 0.71, RMSEP=9.68N SSC: R^2 = 0.81, RMSEP=0.69%	(Nguyen Do Trong et al., 2014a)
	Apple	Bruise detection	PLS	Rp=0.848-0.919, RMSEP=32.4-50.7	(Zhu et al., 2016)
	Apple	Mealiness classification	PCR, PLSR, ANN	Non-mealy: 76% Mealt: 82% Fresh: 88% Semi-mealy: 59%	(Mollazade and Arefi, 2017)
	Apple	Varieties	PLSDA	Classification accuracies=0.994,	(Huang et al., 2020b)
	Apple	Firmness and SSC	CARS, PLSR	Firmness: R^2 = 0.96, $RMSE_{cal}$=0.37N, SSC: R^2 = 0.87, $RMSE_{cal}$=0.71N	(Ma et al., 2021b)
	Apple	Anthocyanins	PLS	Skin: R^2>0.95, Whole flesh R^2 = 0.74	(Ye et al., 2021)

(Continued)

TABLE 2 Continued

Products	Species	Applications	Methods	Accuracy	References
	Peach	Maturity/quality assessment	PLS, PCA, LS-SVM	Firmness: 0.794, SSC: 0.504, Skin lightness: 0.898, Flesh lightness: 0.741	(Cen et al., 2011; Cen et al., 2012a)
	Peach	Tissue structural and biochemical properties	SPA, PCA	Membrane permeability μ_s'=-0.962-0.743	(Sun Y. et al., 2020)
	Peach	Bruise detection	ANOVA, LSD SVM, PLSDA, C-SVC	μ_a=76.25%, μ_s'=76.25%, $\mu_a \times \mu_s$'=84.75%, μ_{eff}=84.5%	(Sun Y. et al., 2021)
	Peach	Firmness and SSC	PLS	Firmness: 0.853, SSC: 0.839	(Huang et al., 2022)
	Peach	pear porosity	ANOVA	760nm: R^2 = 0.66 835nm: R^2 = 0.57	(Joseph et al., 2023)
	Kiwifruit	Firmness, SSC, pH	PLSR	Firmness: R^2 = 0.37, SSC: R^2 = 0.81, pH: R^2 = 0.59	(Ma et al., 2022)
	Pear	Optical property analysis (μ_a, μ_s')	–	μ_a=0.10-0.61cm^{-1} μ_s'=12.5-9.5cm^{-1}	(Hu et al., 2017)
	Banana	Chlorophyll, elasticity, SSC, ripeness	ANN	CH: R=0.9768-0.9807, EL: R=0.9553-0.9759, SSC: R=0.9640-0.9801, RI: classification, accuracy=97.53%	(Adebayo et al., 2016)
	Citrus	Early decay detection	GL, LDA	Classification accuracy=96.1%	(Lorente et al., 2013)
Vegetables	Cucumber	Defect detection	–	–	(Lu et al., 2011)
	Onion	Detecting internal rots	PLSDA	–	(Sun J. et al., 2020)
	Tomato	Maturity classification	PLSDA, SVMDA	Classification accuracy=81.3–96.3%	(Huang and Chen, 2018)
	Tomato	SSC, pH	PLS	SSC: r_p=0.800, pH: r_p=0.819	(Huang et al., 2018a)
	Tomato	Firmness, SSC, pH	PLS	Firmness: R=0.835, SSC: R=0.623, pH: R=0.769	(Huang et al., 2018b)
	Tomato	Firmness, puncture maximum force, slope	PLS	F: 0.859, PMF: 0.917, SL: 0.948	(Huang et al., 2018c)
	Tomato	Maturity stages	SVMDA	Total classification accuracy=98.3%	(Huang et al., 2020a)
	Tomato	Ripeness	PLS-DA	Classification accuracy=88.4%	(Zhu et al., 2015)
Wood	Softwood silver fir	Dry, wet	MC	Dry: μ_a=0.0048mm^{-1}, 0.0042mm^{-1}, μ_s'=1.8mm^{-1}, 13 mm^{-1}, Wet: μ_a=0.0045mm^{-1}, 0.0038mm^{-1}, μ_s'=0.6mm^{-1}, 2.0mm^{-1}	(Kienle et al., 2008)
	Douglas fir	Various densities, grain directions, thicknesses	PCR, PLS	3mm: R=0.953, 5mm: R=0.987	(Ma et al., 2018)
	Five softwood (SW) ten	Classification	PCA, QDA	QDA=94.0%	(Ma et al., 2019a)

(Continued)

TABLE 2 Continued

Products	Species	Applications	Methods	Accuracy	References
	hardwood (HW)				
	Hinoki cypress	Three-dimensional grain angle	GPR, LRA	GPR: R^2 = 0.98, RMSE=2.2° LRA: R^2>0.90, RMSE<3.8°	(Ma et al., 2019b)
	Wood	Tensile strain measurement	PCA, PLSR	R^2 = 0.86, RMSE=279.86	(Ma et al., 2021c)
	Wood	Classification	PCA, SVM	Five-fold cross-validation=98.6%, Test set validation=91.2%	(Ma et al., 2021a)
Animal Husbandry	Cattle	Age	PLS, GA, RLT	ARMSEP=2.0 years, R^2 = 0.63	(Palendeng et al., 2020)
	Rabbit	Early pregnancy diagnosis	PLS-DA, CARS, SPA, SPA, SVM, KNN, Naïve Bayes	Validation set Sensitivity=93.18%, Specificity=94.44%, Accuracy=93.88%, Prediction set Sensitivity=86.96%, Specificity=90.00%, Accuracy=90.69%	(Yuan et al., 2022)

JVSPO, Joint variable selection and preprocessing optimization method; MLD, Modified Lorentzian distribution; GD, 'Golden Delicious'; RD, 'Delicious' (RD) apples; ANOVA, Analysis of variance; LSD, Least significant difference; LDA, Linear discriminant analysis; C-SVC, C-Support Vector Classification algorithm; GL, Gaussian-Lorentzian cross product; r_p, Correlation coefficient of prediction; PCR, Principal component regression; PLS, Partial least squares regression analysis; GPR, Gaussian process regression; RMSE, Root mean square error; LRA, Linear regression analysis; PCA, Principal component analysis; QDA, Quadratic discriminant analysis; PLS, Partial least squares regression; GA, Genetic algorithm; RLT, Repeated learning-training method; ARMSEP, Average root mean square error of prediction; MC, Monte Carlo simulations; PLS-DA, Partial least squares-discriminant analysis; CARS, Competitive adaptive reweighted sampling; SPA, Successive projection algorithm; SVM, Support vector machine; KNN, K-Nearest Neighbor.

can be tested by this method, SRS has excellent potential for future applications in meat quality testing.

4.3 Applications of fruit and vegetable field

SRS is widely applied in fruit inspection, mainly for apples, pears, peaches, kiwifruit, bananas, and citrus. In the detection process of apples, spatially resolved hyperspectral imaging was to measure apple OP and relate them to fruit firmness and SSC, showing that the μ_a and μ_s' data gave the best predictions for the fruit firmness and SSC, with correlation coefficients (r) of 0.82 and 0.80 for firmness, and 0.7 and 0.59 for SSC respectively. This provides a fresh approach to detecting the internal quality of fruits (Qin et al., 2007). Peng and Lu refined the hyperspectral scattering technique for fruit quantity testing by fitting spectral scattering curves at each wavelength with ten different forms of modified Lorentzian distribution functions and comparing the predictions of fruit firmness and SSC by ten modified Lorentzian distribution functions using multiple linear regression and cross-validation methods. The predicted correlation coefficients were 0.894 and 0.883, respectively, which verified the advantages of the technique in fruit quality testing (Peng and Lu, 2008). Lu et al. used the absorption scattering properties of apple tissue to predict bruising of the fruit. The measurement of enhanced scattering properties was found to be feasible for bruise detection in apples (Lu et al., 2010). Huang et al. detected the mealiness of apples, modeled the classification of apple mealiness classes by the partial least squares (PLS) method, and found that the accuracy of establishing a two-level classification was ≥93%. Thus, it validated the advantages of hyperspectral scattering technology in the nondestructive detection of the mealiness of apples (Huang and Lu, 2010). Cen et al. analyzed the physical and structural properties of apple pulp using a newly developed OPA (Cen et al., 2012b) and the correlation coefficients of firmness, r=0.870-0.948, and Young's modulus, r=0.585-0.947, were obtained for Golden Delicious (GD) apples, which demonstrated that spatially resolved techniques can be used to predict internal fruit quality by combining OP (Cen et al., 2013). Mendoza et al. used short-wave NIR spectroscopy and scattering to classify apple quality with accuracies ranging from 87.3-97.6% for firmness and 77.1-92.3% for SSC, which validated the capability of organizing and grading apples by firmness and SSC (Mendoza et al., 2014). Nguyen Do Trong obtained the scattering and absorption coefficients of apple slices air-dried under various conditions pretreated by spatially resolved diffuse reflectance spectroscopy. Finally, it was found that SRS could detect the microstructure and quality relationship of air-dried apple slices without loss (Nguyen Do Trong et al., 2014b). The spatially resolved diffuse reflectance device (Nguyen Do Trong et al., 2013) was used to detect the OP of apples (Nguyen Do Trong et al., 2014a). The μ_a spectrum was found to be superior to μ_s' by comparison, and the coefficients of determination R^2 for firmness and SSC were 0.71 and 0.81, respectively. The results showed that the detection of diffuse reflectance spectra of optical fibers cannot significantly improve the prediction performance of SSC. Still, it can be used to better predict the firmness and SSC of apples by separating the absorption coefficients and reducing the scattering coefficients. Zhu et al. utilized hyperspectral scattering to expected damage to apples

with predictive correlation coefficients R_p=0.848-0.919. The research revealed that hyperspectral scattering can be used to assess the bruise susceptibility of apples, which is beneficial for post-harvest inspection of fruits (Zhu et al., 2016). Mollazade et al. found a way to classify apple fruits using spatial resolving technique, which was verified by 76% and 82% accuracy for non-mealy and mealy apples, respectively (Mollazade and Arefi, 2017). Huang et al. used a multichannel HSI to classify apple varieties with 99.4% accuracy using the best spectral classification. They verified the potential of multichannel hyperspectral imaging systems for apple variety detection (Huang et al., 2020b). Ma and Xia et al. used a multi-fiber, spatially resolved measurement system that assessed the SSC and firmness of apples with an optimal R^2 of 0.97 and 0.96, respectively, validating the technique's ability to detect apple quality in a low-cost and portable method accurately (Ma et al., 2021b). Ye et al. obtained spatially resolved interaction spectra at eight different source-detector distances (SDDs) on the fruit surface and verified that the optimal SD could be selected to detect the extent of red color in the flesh at a specific depth by a model developed for anthocyanin content estimation (Ye et al., 2021).

In the inspection of peaches, Cen et al. (2011; 2012a) measured the absorption and reduced scattering coefficients based on the SR method of HSI to assess peach ripeness and quality, with *r* of 0.749 and 0.504 for firmness and SSC, respectively. The results suggested that spatially resolved techniques had good potential for application. Research by Sun et al. measured the OP of peaches during quality damage, determined the relationship between optical parameters and specific structural and biochemical factors, and found a good correlation at 675 nm (Sun Y. et al., 2020). This study facilitated the early detection of peach diseases. Sun et al. also measured the OP of peaches at different ripeness levels using the SR technique (Sun Y. et al., 2021), with classification accuracies of 85% and 76.25%, respectively, and these results found that this optical property was effective in detecting damage in peaches. Huang et al. evaluated the firmness and SSC of peaches using SRS (Huang et al., 2022) and improved the prediction of peach quality by incorporating spectral disparity techniques, with the best *r* of peach firmness and SSC being 0.853 and 0.839, respectively. Joseph et al. used the SRS technique to study the relationship between peach porosity and light scattering characteristics, and the results showed that the reduced scattering coefficients at 760 nm and 835 nm were linearly correlated with the spatially averaged porosity by R^2 of 0.66 and 0.57, respectively, which verified that the method could realize non-destructive pear porosity assessment (Joseph et al., 2023).

In addition, Ma et al. verified the feasibility of SRS for the detection of kiwifruit quality with coefficients of determination R^2 of 0.81 and 0.59 for SSC and pH, respectively (Ma et al., 2022). The OP of the pear was analyzed by Hu et al. (2017). They measured μ_a between 0.1-0.61 cm^{-1}, while μ_s' decreased with wavelength between 12.5-9.5 cm^{-1}. In this study, it was demonstrated that the OP of pears is associated with the wavelength and that establishing standardized slices of the samples helps to enhance the precision of the measurement of the OP. Adebayo et al. combined OP with chlorophyll, modulus of elasticity, SSC, and banana ripeness to develop predictive models. The correlation coefficients of chlorophyll, elastic modulus, and SSC were 0.9768-0.9807, 0.9553-0.9759, and 0.9640-0.9801, respectively, and the classification accuracy of banana ripeness reached 97.53%. This indicates that bananas with different ripeness levels can be predicted and categorized by OP, which provides a good and effective method for nondestructive testing of banana quality (Adebayo et al., 2016). Lorente et al. predicted early decay in citrus fruits with a classification accuracy of 96.1%, validating that this technique has great potential for grading citrus fruits (Lorente et al., 2013).

In the detection of vegetables, Lu et al. used the spatially resolved technique of hyperspectral imaging to test for defective pickling cucumbers (Lu et al., 2011). They found that effective defect detection could be achieved by enhanced scattering characteristic measurements through analysis of the OP of cucumbers. Sun et al. developed the SR transmission spectroscopy system for detecting internal rot onions, and the presence of high area under curve (AUC) values (0.96 ± 0.02) and Kappa values (0.77 ± 0.05) at the stem end of the onion validated the advantages of the system in detecting onion decay (Sun J. et al., 2020). Huang et al. designed a multichannel SRS detection device and used it to detect firmness, SSC, pH with correlation coefficients of 0.835, 0.623, and 0.769, respectively (Huang et al., 2018a; Huang et al., 2018b), The classification accuracy in tomato maturity assessment was able to reach 98.3% (Huang et al., 2020a; Huang et al., 2020b), which verified that OP based on SRS can reasonably predict the quality of tomato.

Table 2 shows the details of the studies on the detection of product quality of fruits and vegetables. It shows that SRS has been widely used in the field of fruits and vegetables, mainly for the detection of quality characteristics such as firmness, pH, SSC, maturity, mealiness and bruise, as well as the biochemical properties of the internal tissues. In the future, the application of SRS in fruit and vegetable detection will be more mature, the types of detection will be more abundant, and the accuracy will be higher.

4.4 Applications of forestry field

In the field of forestry industry, SRS is mainly used in the detection of wood in recent years. Kienle et al. used a spatially and time resolved approach to study the mechanism of light propagation in dry and moist softwoods and put forward a theoretical model for the description of light propagation in wood, which is supported by the microstructure of softwood (Kienle et al., 2008). Ma et al. used spatially resolved hyperspectral detection to examine the OP, grain direction, and thickness of Douglas-fir at different densities. Correlation coefficients for 3 mm and 5 mm samples were 0.953 and 0.987, respectively (Ma et al., 2018). Meanwhile, the device classified softwoods and hardwoods with an accuracy of 94.1%, which shows that SRS is highly predictive in wood inspection (Ma et al., 2019a). In addition, the SRS device was optimized to achieve 91.2% accuracy in the test set of 15 wood classifications (Ma et al., 2021a). Moreover, the R^2 of the tensile strain of the wood was measured to be 0.86 using the optimized equipment, which verified the suitability of SRS for the detection of wood (Ma et al., 2021c).

Table 2 summarizes in detail the current status of product quality testing in forestry. It can be found that SRS detection technology is emerging in the application of forestry quality testing, and with the progress of technology, this method will be more widely used in forestry-related testing.

4.5 Applications of animal husbandry field

In the field of animal husbandry, Palendeng et al. used a spatially resolved method to detect the age of cattle (Palendeng et al., 2020). The feasibility of the SR technique for estimating the age of cattle was validated by using the SR diffuse reflectance spectrometer based on a fiber optic probe to collect skin samples from the neck of the cattle and assessing the age of the cattle by the developed PLS model with the lowest average root mean square error of prediction (ARMSEP) of 2.0 years and $R^2 = 0.63$. Yuan et al. used SRS to diagnose the possibility of pregnancy in female rabbits by collecting spectral information at different distances with a movable distance-type detection fiber (Yuan et al., 2022). The results showed that the SRS detection method can distinguish whether a female rabbit is pregnant or not, and the accuracy of the validation set can reach 90.69%.

From Table 2, it can be found that the application of SRS technology in animal husbandry-related fields is still rare, and it is currently only applied to a few animal husbandry tests, mainly for age and pregnancy detection of animals. However, this technique shows a strong predictive ability in livestock detection. Therefore, the method is expected to be widely applied to the detection of other characteristics of animal husbandry and more livestock animals in future development.

Generally, SRS technology has been widely used in the field of agriculture. In the field of dairy this technique is mainly applied to the detection of protein, fat and other nutrients in milk with better predicted results. But at present the technique is less used for the detection of other types of dairy products such as goat milk, camel milk, etc., and some dairy products such as milk powder, cream, cheese, etc. Hence there is a good prospect for development in this field. In the detection of meat products, researchers mainly focus on the detection of fresh beef and pork, and the qualities detected are meat tenderness, myoglobin and moisture content. Nevertheless, the quality of some meat products such as jerky, dried meat, bacon, sausage, etc. was less tested. In the future, other types of meat can also be detected, such as lamb, fish, shrimp, etc., through the detection of its nutrient content to predict the quality of meat products, which is conducive to providing human beings with more healthy and nutritious food. SRS has been most used and developed for fruits and vegetables. Currently, the fruits and vegetables inspected include apples, pears, kiwifruits, bananas, citrus, cucumbers, onions, and tomatoes. The main detection of their SSC, hardness, pH, damage, ripeness, chlorophyll content, etc. Thus, SRS is expected to achieve more efficient and accurate quality detection in this field. In the field of forestry, researchers mainly apply SRS to the classification, moisture, texture, and thickness detection of wood. In future development, it is expected to realize the detection of hardness, oiliness, density and damage of wood, which has a lot of space for development. In the field of animal husbandry, the current research is mainly focused on the age of cows and the pregnancy of rabbits, but in the future, it is expected to detect more animals and their health level. Overall, SRS has been more widely used in the field of agriculture at present but still has a lot of advantages for development. Since this technique can not only directly analyze the correlation through spatial spectral information, but also extract specific optical properties to further explore the relationship between the quality of agricultural products and OP. Therefore, it is expected that this technology will have much more effective application potential in future quality detection in the field of agriculture.

5 Challenges and future trends

SRS has been widely used in agriculture so far due to its stable performance, low cost, ease of use, and continuous algorithmic improvement. Importantly, this is mainly because that detection method can well reflect the characteristics of agricultural products. Although five different types of SRS, including single fiber, fiber array type, CCD line scan type, hyperspectral line-scan, and multi-channel hyperspectral imaging detection system, are relatively widely used in agricultural products for quality inspection currently, this technology still faces many challenges and difficulties.

The challenges are mainly in the SRS devices and calculation methods. In terms of devices, for example, there is no specific standard for the selection of light sources, and the selection of high-power light sources can easily damage the external and internal tissue structure of organisms. In contrast, the selection of low-power light sources has limited detection distance and cannot collect satisfactory distant spectral information. To meet the requirement of detection, the light source should satisfy the appropriate intensity meanwhile its diameter is often small enough, especially for small samples such as corn kernels, wheat seeds, cherries, grapes, and other agricultural products, so that it can be equivalent to a point light source and reduce the error of solving the OP of the calculation. The practical arrangement and selection of optical fiber is also a problematic issue in the device. For single fiber and fiber array detection devices, the selection of optical fiber is significant, which not only requires the fiber to be as small as possible but also the arrangement of the detection distance as accurately as possible. Besides, for irregular detection objects, the detection fibers often cannot fit closely due to the curvature of the sample surface. Although Huang et al. designed a multi-channel detection device, the approach is only suitable for larger objects with micro-curvature. Some irregular-shaped and curvature-changed objects or smaller objects still cannot be detected satisfactorily. Therefore, how to design a detection fiber that can meet irregular objects is still an essential and inevitable challenge.

In addition, the stability and precision of the mechanical device enable more accurate acquisition of spatially resolved spectral information. Therefore, the design of the mechanics of the SRS is of great significance. During the spectra acquisition process, the detection device needs to hold the fiber and the sample firmly in place. Manual detection often lacks accuracy, which is prone to

jitter, and has many instabilities. These can undoubtedly have a negative impact on the spectral quality. As a result, more stable mechanical devices are needed to replace manual fixation to improve the stability of the detection system. However, the mechanical device has different requirements for the detection fiber, light source, and sample. As for optical fiber detection, it is not only required that the fiber closely fits the sample but also that the fiber is moved in a more precise position. However, because of the irregularity of the measured object and the curvature of the surface, it is easy for the optical fibers to move without close contact, and can also damage the surface of the agricultural product if it is moved too aggressively. In the case of the light source, it is necessary that the light source is also close-fitting the measured object surface to avoid too much diffuse light on the spectral information. As for the sample, the mechanical device should be fixed steadily so that the collected sample cannot move easily, and it should not be fixed too tightly to avoid damage or deformation to the sample. So, there are many difficulties in the design of the mechanical device.

Moreover, the detection accuracy has a great impact on the subsequent analysis of the spectra. Trying to minimize the impact of some controllable factors on the accuracy is beneficial to improving the spectral detection quality. The detection accuracy is affected by various factors, which are reflected in all aspects of the detection device. For example, the resolution of the detection instrument, the stability of the light source, the loss of the detection fiber, and the stability of the mechanical device might have a negative impact on the accuracy. Therefore, the designed SRS equipment requires calibration to guarantee the stability and accuracy of the equipment. Last but not least, the design cost of SRS is also a problem because the manufacturing and maintenance costs of SRS detection devices are very high and usually require the use of expensive optical and mechanical components. How to improve the detection accuracy of SRS while reducing the cost is also a demanding challenge to be solved.

At the present stage, direct analysis and OP are mainly used to deal with spatially resolved spectra in terms of computational methods. The direct analysis method is simple and efficient, but this method is less used. Therefore, it is a promising trend and a good development direction to study the simple and efficient direct analysis method. In addition, the OP method is complex, but the prediction accuracy is relatively high. The current OP methods are based on the diffuse equation theory to separate out the optical characteristics. In the process of extracting the OP, because of the complexity of the diffuse equation, the solved values are often not accurate enough and the computation is huge. The future development of simpler and more accurate optical equations to extract OP based on the current research is also an emerging research prevailing trend.

6 Conclusion

Agricultural products, including dairy, meat, fruit and vegetable, forestry products, and animal husbandry products, are of great importance to people's daily lives, depending on their external and internal quality. Compared with traditional detection methods, SRS not only provides more spatial information but also separates out optical properties, so it can be widely used in the field of agriculture. SRS detection systems, including single fiber, fiber array type, CCD line scan type, hyperspectral line-scan, and multi-channel hyperspectral imaging detection system, have been increasingly used for inspecting quality in replacement of manual grading as they can provide a simple structure, easy to operate, low cost and non-destructive assessment. With the continuous development of this technology, many successful applications have proved that SRS detection systems are powerful and scientific tools for stable and accurate quality inspection of agricultural products. This paper reviews the principles, development, and applications of five various SRS detection systems for agricultural product quality inspections. Despite the problems and challenges of this technique, it promises to achieve online detection with a more simple, portable, and easy to operate configuration for further widespread application in quality inspection of agricultural fields.

Author contributions

YX: Writing – original draft, Writing – review & editing, Conceptualization, Methodology, Funding acquisition, Supervision. WXL: Writing – original draft, Writing – review & editing, Conceptualization, Investigation, Data curation, Methodology. JM: Data curation, Software, Writing – original draft. JH: Formal Analysis, Writing – original draft. WBL: Writing – review & editing, Resources. JK: Writing – review & editing, Funding acquisition, Project administration. BL: Writing – review & editing, Resources. HZ: Writing – review & editing, Resources, Supervision. WT: Writing – review & editing, Resources, Supervision.

Funding

The author(s) declare financial support was received for the research, authorship, and/or publication of this article. This work was supported by the Natural Science Basic Research Program of Shaanxi (2022JQ-181), the National Natural Science Foundation of China (62203285) and the Xi'an Science and Technology Plan Project (23NYGG0070).

Conflict of interest

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

Publisher's note

References

Adebayo, S. E., Hashim, N., Abdan, K., Hanafi, M., and Mollazade, K. (2016). Prediction of quality attributes and ripeness classification of bananas using optical properties. *Scientia Hortic.* 212, 171–182. doi: 10.1016/j.scienta.2016.09.045

Audu, J., and Aremu, A. K. (2021). Development, evaluation, and optimization of an automated device for quality detection and separation of cowpea seeds. *Artif. Intell. Agric.* 5, 240–251. doi: 10.1016/j.aiia.2021.10.003

Bao, C., Zhang, H., Li, Q., Zhou, S., Zhang, H., Deng, K., et al. (2021). Spatially-resolved electronic structure of stripe domains in IrTe2 through electronic structure microscopy. *Commun. Phys.* 4, 229. doi: 10.1038/s42005-021-00733-x

Bogomolov, A., Belikova, V., Galyanin, V., Melenteva, A., and Meyer, H. (2017). Reference-free spectroscopic determination of fat and protein in milk in the visible and near infrared region below 1000nm using spatially resolved diffuse reflectance fiber probe. *Talanta* 167, 563–572. doi: 10.1016/j.talanta.2017.02.047

Bridger, K. G., Roccabruna, J. R., and Baran, T. M. (2021). Optical property recovery with spatially-resolved diffuse reflectance at short source-detector separations using a compact fiber-optic probe. *Biomed. Opt. Express* 12, 7388–7404. doi: 10.1364/BOE.443332

Cen, H., and Lu, R. (2009). Quantification of the optical properties of two-layer turbid materials using a hyperspectral imaging-based spatially-resolved technique. *Appl. Opt*. 48, 5612–5623. doi: 10.1364/ao.48.005612

Cen, H., and Lu, R. (2010). Optimization of the hyperspectral imaging-based spatially-resolved system for measuring the optical properties of biological materials. *Opt. Express* 18, 17412–17432. doi: 10.1364/OE.18.017412

Cen, H., Lu, R., A. Mendoza, F., and P. Ariana, D. (2012a). Assessing multiple quality attributes of peaches using optical absorption and scattering properties. *Trans. ASABE* 55, 647–657. doi: 10.13031/2013.41366

Cen, H., Lu, R., and Dolan, K. (2010). Optimization of inverse algorithm for estimating the optical properties of biological materials using spatially-resolved diffuse reflectance. *Inverse Problems Sci. Eng.* 18, 853–872. doi: 10.1080/17415977.2010.492516

Cen, H., Lu, R., and Mendoza, F. A. (2012b). Analysis of absorption and scattering spectra for assessing the internal quality of apple fruit. *Acta Hortic.* 945, 181–188. doi: 10.17660/ActaHortic.2012.945.24

Cen, H., Lu, R., Mendoza, F. A., and Ariana, D. P. (2011). Peach maturity/quality assessment using hyperspectral imaging-based spatially resolved technique, in Proceedings of SPIE 8027, Sensing for Agriculture and Food Quality and Safety III, (Bellingham, USA: 80270L, Society of Photo-optical Instrumentation Engineers). doi: 10.1117/12.883573

Cen, H., Lu, R., Mendoza, F., and Beaudry, R. M. (2013). Relationship of the optical absorption and scattering properties with mechanical and structural properties of apple tissue. *Postharvest Biol. Technol.* 85, 30–38. doi: 10.1016/j.postharvbio.2013.04.014

Chen, Y., Xu, Z., Tang, W., Hu, M., Tang, D., Zhai, G., et al. (2021). Identification of various food residuals on denim based on hyperspectral imaging system and combination optimal strategy. *Artif. Intell. Agric.* 5, 125–132. doi: 10.1016/j.aiia.2021.06.001

Cheong, W. F., Prahl, S. A., and Welch, A.J.J.I.J.O.Q.E. (1990). A review of the optical properties of biological tissues. *IEEE J. Quantum Electron.* 26, 2166–2185. doi: 10.1109/3.64354

Chong, K. C., and Pramanik, M. (2023). Physics-guided neural network for tissue optical properties estimation. *Biomed. Opt. Express* 14, 2576–2590. doi: 10.1364/BOE.487179

Colas, V., Amouroux, M., Perrin-Mozet, C., Daul, C., and Blondel, W. (2023). Photometric and Monte-Carlo modeling unified approach for the calculation of spatially-resolved correction coefficients linking simulated and experimental diffuse reflectance spectra. *Opt. Express* 31, 25954–25969. doi: 10.1364/OE.491921

Comerford, J. M., Negus, J., Barrows, R. S., Wylezalek, D., Greene, J. E., Müller-Sánchez, F., et al. (2022). Toward a more complete optical census of active galactic nuclei via spatially resolved spectroscopy. *Astrophys J.* 927, 23. doi: 10.3847/1538-4357/ac496a

Cubeddu, R., D'andrea, C., Pifferi, A., Taroni, P., Torricelli, A., Valentini, G., et al. (2001). Nondestructive quantification of chemical and physical properties of fruits by time-resolved reflectance spectroscopy in the wavelength range of 650–1000 nm. *Appl. Opt.* 40, 538–543. doi: 10.1364/ao.40.000538

Dam, J. S., Pedersen, C. B., Dalgaard, T., Fabricius, P. E., Aruna, P., and Andersson-Engels, S. (2001). Fiber-optic probe for noninvasive real-time determination of tissue optical properties at multiple wavelengths. *Appl. Opt.* 40, 1155–1164. doi: 10.1364/ao.40.001155

De Man, A., Uyttersprot, J. S., Chavez, P. F., Vandenbroucke, F., Bovart, F., and De Beer, T. (2023). The application of Near-Infrared Spatially Resolved Spectroscopy in scope of achieving continuous real-time quality monitoring and control of tablets with challenging dimensions. *Int. J. Pharm.* 641, 123064. doi: 10.1016/j.ijpharm.2023.123064

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

Doornbos, R. M., Lang, R., Aalders, M. C., Cross, F. W., and Sterenborg, H. J. (1999). The determination of in *vivo* human tissue optical properties and absolute chromophore concentrations using spatially resolved steady-state diffuse reflectance spectroscopy. *Phys. Med. Biol.* 44, 967–981. doi: 10.1088/0031-9155/44/4/012

Farrell, T. J., Patterson, M. S., and Wilson, B. (1992). A diffusion theory model of spatially resolvedsteady-state diffuse reflectance for the non-invasive determination of tissue optical properties in *vivo*. *Med. Phys.* 19, 879–888. doi: 10.1118/1.596777

Guo, Z., Zhang, Y., Wang, J., Liu, Y., Jayan, H., El-Seedi, H. R., et al. (2023). Detection model transfer of apple soluble solids content based on NIR spectroscopy and deep learning. *Comput. Electron. Agric.* 212, 108127. doi: 10.1016/j.compag.2023.108127

Hank, P., Blum, C., Liemert, A., Geiger, S., and Kienle, A. (2023). Analytical solution for the single scattered radiance of two-layered turbid media in the spatial frequency domain. Part 2: Vector radiative transfer equation. *Optics Commun.* 535, 129354. doi: 10.1016/j.optcom.2023.129354

Haskell, R. C., Svaasand, L. O., Tsay, T. T., Feng, T. C., Mcadams, M. S., and Tromberg, B. J. (1994). Boundary conditions for the diffusion equation in radiative transfer. *Opt Soc. Am. A Opt Image Sci. Vis.* 11, 2727–2741. doi: 10.1364/josaa.11.002727

Hu, D., Fu, X., and Ying, Y. (2017). Characterizing pear tissue with optical absorption and scattering properties using spatially-resolved diffuse reflectance. *J. Food Meas. Charact.* 11, 930–936. doi: 10.1007/s11694-017-9465-x

Hu, D., Lu, R., and Ying, Y. (2018). A two-step parameter optimization algorithm for improving estimation of optical properties using spatial frequency domain imaging. *J. Quant. Spectrosc. Radiat. Transf.* 207, 32–40. doi: 10.1016/j.jqsrt.2017.12.022

Hu, D., Lu, R., and Ying, Y. (2020a). Spatial-frequency domain imaging coupled with frequency optimization for estimating optical properties of two-layered food and agricultural products. *J. Food Eng.* 277, 109909. doi: 10.1016/j.jfoodeng.2020.109909

Hu, D., Sun, T., Yao, L., Yang, Z., Wang, A., and Ying, Y. (2020b). Monte Carlo: A flexible and accurate technique for modeling light transport in food and agricultural products. *Trends Food Sci. Technol.* 102, 280–290. doi: 10.1016/j.tifs.2020.05.006

Huang, R. L. Q. C., and Chen, K.-J. (2018). Tomato maturity classification based on spatially resolved spectra. *Spectrosc. Spectral Anal.* 38, 2183–2188. doi: 10.3964/j.issn.1000-0593(2018)07-2183-06

Huang, M., and Lu, R. (2010). Apple mealiness detection using hyperspectral scattering technique. *Postharvest Biol. Technol.* 58, 168–175. doi: 10.1016/j.postharvbio.2010.08.002

Huang, Y., Lu, R., and Chen, K. (2017). Development of a multichannel hyperspectral imaging probe for property and quality assessment of horticultural products. *Postharvest Biol. Technol.* 133, 88–97. doi: 10.1016/j.postharvbio.2017.07.009

Huang, Y., Lu, R., and Chen, K. (2018a). Assessment of tomato soluble solids content and pH by spatially-resolved and conventional Vis/NIR spectroscopy. *J. Food Eng.* 236, 19–28. doi: 10.1016/j.jfoodeng.2018.05.008

Huang, Y., Lu, R., Hu, D., and Chen, K. (2018b). Quality assessment of tomato fruit by optical absorption and scattering properties. *Postharvest Biol. Technol.* 143, 78–85. doi: 10.1016/j.postharvbio.2018.04.016

Huang, Y., Lu, R., Xu, Y., and Chen, K. (2018c). Prediction of tomato firmness using spatially-resolved spectroscopy. *Postharvest Biol. Technol.* 140, 18–26. doi: 10.1016/j.postharvbio.2018.02.008

Huang, Y., Si, W., Chen, K., and Sun, Y. (2020a). Assessment of tomato maturity in different layers by spatially resolved spectroscopy. *Sens. (Basel)* 20, 7229. doi: 10.3390/s20247229

Huang, Y., Xiong, J., Jiang, X., Chen, K., and Hu, D. (2022). Assessment of firmness and soluble solids content of peaches by spatially resolved spectroscopy with a spectral difference technique. *Comput. Electron. Agric.* 200, 107212. doi: 10.1016/j.compag.2022.107212

Huang, Y., Yang, Y., Sun, Y., Zhou, H., and Chen, K. (2020b). Identification of apple varieties using a multichannel hyperspectral imaging system. *Sens. (Basel)* 20, 5120. doi: 10.3390/s20185120

Igne, B., Talwar, S., Feng, H., Drennen, J. K., and Anderson, C. A. (2015). Near-infrared spatially resolved spectroscopy for tablet quality determination. *J. Pharm. Sci.* 104, 4074–4081. doi: 10.1002/jps.24618

Joseph, M., Van Cauteren, H., Postelmans, A., Nugraha, B., Verreydt, C., Verboven, P., et al. (2023). Porosity quantification in pear fruit with X-ray CT and spatially resolved spectroscopy. *Postharvest Biol. Technol.* 204, 112455. doi: 10.1016/j.postharvbio.2023.112455

Kalinin, A. V., Krasheninnikov, V. N., and Krivtsun, V. M. (2013). Short-wave near infrared spectrometry of back scattering and transmission of light by milk for multi-component analysis. *J. Near Infrared Spectrosc.* 21, 35–41. doi: 10.1255/jnirs.1034

Katsumata, K. (2022). A theory of reflection and refraction of light by a metal. *Optical Rev.* 29, 159–171. doi: 10.1007/s10043-022-00732-5

Kienle, A., D'andrea, C., Foschum, F., Taroni, P., and Pifferi, A. (2008). Light propagation in dry and wet softwood. *Opt. Express* 16, 9895–9906. doi: 10.1364/oe.16.009895

Kienle, A., Lilge, L., Patterson, M. S., Hibst, R., Steiner, R., and Wilson, B. C. (1996). Spatially resolved absolute diffuse reflectance measurements for noninvasive determination of the optical scattering and absorption coefficients of biological tissue. *Appl. Opt.* 35, 2304–2314. doi: 10.1364/ao.35.002304

Kienle, A., and Patterson, M. S. (1997). Improved solutions of the steady-state and the time-resolved diffusion equations for reflectance from a semi-infinite turbid medium. *Opt Soc. Am. A Opt. Image Sci. Vis.* 14, 246–254. doi: 10.1364/josaa.14.000246

Kienle, A., Patterson, M. S., Dögnitz, N., Bays, R., Wagnivres, G., and Van Den Bergh, H. (1998). Noninvasive determination of the optical properties of two-layered turbid media. *Appl. Opt.* 37, 779–791. doi: 10.1364/ao.37.000779

Kubelka, P. (1948). New contributions to the optics of intensely light-scattering materials. Part I. *J. Optical Soc. America* 38, 448–457. doi: 10.1364/JOSA.38.000448

Lan, Q., Mcclarren, R. G., and Vishwanath, K. (2023). Neural network-based inverse model for diffuse reflectance spectroscopy. *Biomed. Opt. Express* 14, 4725–4738. doi: 10.1364/BOE.490164

Lee, J. H., Kim, S., and Kim, Y. T. (2004). Finite element method for diffusive light propagations in index-mismatched media. *Opt. Express* 12 8, 1727–1740. doi: 10.1364/opex.12.001727

Li, M., Gu, J., Zhang, D., Gao, Q., and Li, B. (2021). Equivalence ratio measurements in CH4/air gases based on the spatial distribution of the emission intensity of femtosecond laser-induced filament. *Processes* 9, 2022. doi: 10.3390/pr9112022

Liu, G., Asif, M., Mohri, K., Schulz, C., Dreier, T., Endres, T., et al. (2022). *In situ* measurement of gas-borne silicon nanoparticle volume fraction and temperature by spatially and spectrally line-resolved attenuation and emission imaging. *Powder Technol.* 396, 535–541. doi: 10.1016/j.powtec.2021.11.017

Lorente, D., Zude, M., Regen, C., Palou, L., Gómez-Sanchis, J., and Blasco, J. (2013). Early decay detection in citrus fruit using laser-light backscattering imaging. *Postharvest Biol. Technol.* 86, 424–430. doi: 10.1016/j.postharvbio.2013.07.021

Lu, R., Ariana, D. P., and Cen, H. (2011). Optical absorption and scattering properties of normal and defective pickling cucumbers for 700–1000 nm. *Sens. Instrument. Food Qual. Saf.* 5, 51–56. doi: 10.1007/s11694-011-9108-6

Lu, R., Cen, H., Huang, M., and Ariana, D.P.J.T.O.T.A. (2010). Spectral absorption and scattering properties of normal and bruised apple tissue. *Trans. ASABE* 53, 263–269. doi: 10.13031/2013.29491

Lu, R., Van Beers, R., Saeys, W., Li, C., and Cen, H. (2020). Measurement of optical properties of fruits and vegetables: A review. *Postharvest Biol. Technol.* 159, 111003. doi: 10.1016/j.postharvbio.2019.111003

Ma, T., Inagaki, T., Ban, M., and Tsuchikawa, S. (2019a). Rapid identification of wood species by near-infrared spatially resolved spectroscopy (NIR-SRS) based on hyperspectral imaging (HSI). *Holzforschung* 73, 323–330. doi: 10.1515/hf-2018-0128

Ma, T., Inagaki, T., and Tsuchikawa, S. (2019b). Three-dimensional grain angle measurement of softwood (Hinoki cypress) using near infrared spatially and spectrally resolved imaging (NIR-SSRI). *Holzforschung* 73, 817–826. doi: 10.1515/hf-2018-0273

Ma, T., Inagaki, T., and Tsuchikawa, S. (2021a). Demonstration of the applicability of visible and near-infrared spatially resolved spectroscopy for rapid and nondestructive wood classification. *Holzforschung* 75, 419–427. doi: 10.1515/hf-2020-0074

Ma, T., Inagaki, T., Yoshida, M., Ichino, M., and Tsuchikawa, S. (2021c). Measuring the tensile strain of wood by visible and near-infrared spatially resolved spectroscopy. *Cellulose* 28, 10787–10801. doi: 10.1007/s10570-021-04239-1

Ma, T., Schajer, G., Inagaki, T., Pirouz, Z., and Tsuchikawa, S. (2018). Optical characteristics of Douglas fir at various densities, grain directions and thicknesses investigated by near-infrared spatially resolved spectroscopy (NIR-SRS). *Holzforschung* 72, 789–796. doi: 10.1515/hf-2017-0213

Ma, T., Xia, Y., Inagaki, T., and Tsuchikawa, S. (2021b). Rapid and nondestructive evaluation of soluble solids content (SSC) and firmness in apple using Vis–NIR spatially resolved spectroscopy. *Postharvest Biol. Technol.* 173, 111417. doi: 10.1016/j.postharvbio.2020.111417

Ma, T., Zhao, J., Inagaki, T., Su, Y., and Tsuchikawa, S. (2022). Rapid and nondestructive prediction of firmness, soluble solids content, and pH in kiwifruit using Vis–NIR spatially resolved spectroscopy. *Postharvest Biol. Technol.* 186, 111841. doi: 10.1016/j.postharvbio.2022.111841

Martelli, F., Tommasi, F., Fini, L., Cortese, L., Sassaroli, A., and Cavalieri, S. (2021). Invariance properties of exact solutions of the radiative transfer equation. *J. Quant. Spectrosc. Radiat. Transf.* 276, 107887. doi: 10.1016/j.jqsrt.2021.107887

Mei, M., and Li, J. (2023). An overview on optical non-destructive detection of bruises in fruit: Technology, method, application, challenge and trend. *Comput. Electron. Agric.* 213, 108195. doi: 10.1016/j.compag.2023.108195

Mendoza, F., Lu, R., Ariana, D., Cen, H., and Bailey, B. (2011). Integrated spectral and image analysis of hyperspectral scattering data for prediction of apple fruit firmness and soluble solids content. *Postharvest Biol. Technol.* 62, 149–160. doi: 10.1016/j.postharvbio.2011.05.009

Mendoza, F., Lu, R., and Cen, H. (2014). Grading of apples based on firmness and soluble solids content using Vis/SWNIR spectroscopy and spectral scattering techniques. *J. Food Eng.* 125, 59–68. doi: 10.1016/j.jfoodeng.2013.10.022

Mishra, P., Nordon, A., Mohd Asaari, M. S., Lian, G., and Redfern, S. (2019). Fusing spectral and textural information in near-infrared hyperspectral imaging to improve green tea classification modelling. *J. Food Eng.* 249, 40–47. doi: 10.1016/j.jfoodeng.2019.01.009

Mohd Ali, M., Hashim, N., Bejo, S. K., Jahari, M., and Shahabudin, N. A. (2023). Innovative non-destructive technologies for quality monitoring of pineapples: Recent advances and applications. *Trends Food Sci. Technol.* 133, 176–188. doi: 10.1016/j.tifs.2023.02.005

Mollazade, K., and Arefi, A. (2017). Optical analysis using monochromatic imaging-based spatially-resolved technique capable of detecting mealiness in apple fruit. *Scientia Hortic.* 225, 589–598. doi: 10.1016/j.scienta.2017.08.005

Morimoto, K., Iguchi, A., and Tsuji, Y. (2020). Propagation operator based boundary condition for finite element analysis. *IEEE Photonics J.* 12, 1–13. doi: 10.1109/JPHOT.2020.3015498

Nguyen Do Trong, N., Erkinbaev, C., Nicolaï, B. M., Saeys, W., Tsuta, M., and De Baerdemaeker, J. (2013). Spatially resolved spectroscopy for nondestructive quality measurements of Braeburn apples cultivated in sub-fertilization condition, in Proceedings of SPIE 8881, Sensing Technologies for Biomaterial, Food, and Agriculture. (Bellingham, USA: 88810L, Society of Photo-optical Instrumentation Engineers). doi: 10.1117/12.2030407

Nguyen Do Trong, N., Erkinbaev, C., Tsuta, M., De Baerdemaeker, J., Nicolaï, B., and Saeys, W. (2014a). Spatially resolved diffuse reflectance in the visible and near-infrared wavelength range for non-destructive quality assessment of 'Braeburn' apples. *Postharvest Biol. Technol.* 91, 39–48. doi: 10.1016/j.postharvbio.2013.12.004

Nguyen Do Trong, N., Rizzolo, A., Herremans, E., Vanoli, M., Cortellino, G., Erkinbaev, C., et al. (2014b). Optical properties–microstructure–texture relationships of dried apple slices: Spatially resolved diffuse reflectance spectroscopy as a novel technique for analysis and process control. *Innovative Food Sci. Emerg. Technol.* 21, 160–168. doi: 10.1016/j.ifset.2013.09.014

Nguyen Do Trong, N., Watté, R., Aernouts, B., Herremans, E., Verhoelst, E., Tsuta, M., et al. (2011). *Spatially Resolved Spectroscopy for Non-destructive Quality Inspection of Foods* (Louisville, Kentucky: ASABE), 1111381. doi: 10.13031/2013.37805

Nichols, M. G., Hull, E. L., and Foster, T. H. (1997). Design and testing of a white-light, steady-state diffuse reflectance spectrometer for determination of optical properties of highly scattering systems. *Appl. Opt.* 36, 93–104. doi: 10.1364/ao.36.000093

Niwayama, M., and Unno, N. (2021). Tissue oximeter with selectable measurement depth using spatially resolved near-infrared spectroscopy. *Sensors* 21, 5573. doi: 10.3390/s21165573

Palendeng, M. E., Alvarenga, T. I. R. C., Fowler, S., Hopkins, D. L., Mcgilchrist, P., and Thennadil, S. N. (2020). Estimation of chronological age of cattle using spatially resolved diffuse reflectance measurements of hide. *IEEE Sens. J.* 20, 8673–8682. doi: 10.1109/jsen.2020.2983455

Peng, Y., and Lu, R. (2008). Analysis of spatially resolved hyperspectral scattering images for assessing apple fruit firmness and soluble solids content. *Postharvest Biol. Technol.* 48, 52–62. doi: 10.1016/j.postharvbio.2007.09.019

Qin, J., and Lu, R. (2007). Measurement of the absorption and scattering properties of turbid liquid foods using hyperspectral imaging. *Appl. Spectrosc.* 61, 388–396. doi: 10.1366/000370207780466190

Qin, J., Lu, R., and Peng, Y. (2007). Internal quality evaluation of apples using spectral absorption and scattering properties. *SPIE Optics East* 6761, 67610M. doi: 10.1117/12.751937

Qin, J., Lu, R., and Peng, Y. (2009). Prediction of apple internal quality using spectral absorption and scattering properties. *Am. Soc. Agric. Biol. Engineers* 52, 499–507. doi: 10.13031/2013.26807

Rejeb, A., Rejeb, K., Zailani, S., Keogh, J. G., and Appolloni, A. (2022). Examining the interplay between artificial intelligence and the agri-food industry. *Artif. Intell. Agric.* 6, 111–128. doi: 10.1016/j.aiia.2022.08.002

Reynolds, L., Johnson, C., and Ishimaru, A. (1976). Diffuse reflectance from a finite blood medium: applications to the modeling of fiber optic catheters. *Appl. Optics* 15, 2059–2067. doi: 10.1364/AO.15.002059

Ryo, M. (2022). Explainable artificial intelligence and interpretable machine learning for agricultural data analysis. *Artif. Intell. Agric.* 6, 257–265. doi: 10.1016/j.aiia.2022.11.003

Sassaroli, A., Tommasi, F., Cavalieri, S., Fini, L., Liemert, A., Kienle, A., et al. (2022). Two-step verification method for Monte Carlo codes in biomedical optics applications. *J. Biomed. Optics* 27, 83018. doi: 10.1117/1.JBO.27.8.083018

Sun, C., Aernouts, B., Van Beers, R., and Saeys, W. (2021). Simulation of light propagation in citrus fruit using monte carlo multi-layered (MCML) method. *J. Food Eng.* 291, 110225. doi: 10.1016/j.jfoodeng.2020.110225

Sun, B., Gao, C., and Spurr, R. (2022). Scalar thermal radiation using the adding-doubling method. *Opt. Express* 30, 30075–30097. doi: 10.1364/OE.462580

Sun, Y., Huang, Y., Pan, L., and Wang, X. (2021). Evaluation of the changes in optical properties of peaches with different maturity levels during bruising. *Foods* 10, 388. doi: 10.3390/foods10020388

Sun, J., Künnemeyer, R., Mcglone, A., Tomer, N., and Sharrock, K. (2020). A spatially resolved transmittance spectroscopy system for detecting internal rots in onions. *Postharvest Biol. Technol.* 163, 111141. doi: 10.1016/j.postharvbio.2020.111141

Sun, Y., Lu, R., Pan, L., Wang, X., and Tu, K. (2020). Assessment of the optical properties of peaches with fungal infection using spatially-resolved diffuse reflectance technique and their relationships with tissue structural and biochemical properties. *Food Chem.* 321, 126704. doi: 10.1016/j.foodchem.2020.126704

Sun, Z., Xie, L., Hu, D., and Ying, Y. (2021). An artificial neural network model for accurate and efficient optical property mapping from spatial-frequency domain images. *Comput. Electron. Agric.* 188, 106340. doi: 10.1016/j.compag.2021.106340

Tarasov, A. P., Persheyev, S., and Rogatkin, D. A. (2021). Analysis of the applicability of the classical probabilistic parameters of the Monte Carlo algorithm for problems of light transport in turbid biological media with continuous absorption and discrete scattering. *Quantum Electron.* 51, 408–414. doi: 10.1070/QEL17535

Tian, X., Liu, X., He, X., Zhang, C., Li, J., and Huang, W. (2023). Detection of early bruises on apples using hyperspectral reflectance imaging coupled with optimal wavelengths selection and improved watershed segmentation algorithm. *J. Sci. Food Agric.* 103, 6689–6705. doi: 10.1002/jsfa.12764

Tian, X., Zhang, C., Li, J., Fan, S., Yang, Y., and Huang, W. (2021). Detection of early decay on citrus using LW-NIR hyperspectral reflectance imaging coupled with two-band ratio and improved watershed segmentation algorithm. *Food Chem.* 360, 130077. doi: 10.1016/j.foodchem.2021.130077

Vanolia, M., Van Beers, R., and Sadarc, N. (2020). Time- and spatially-resolved spectroscopy to determine the bulk optical properties of 'Braeburn' apples after ripening in shelf life. *Postharvest Biol. Technol.* 168, 111233. doi: 10.1016/j.postharvbio.2020.111233

Vasudevan, V., and Narayanan Unni, S.J.I.J.F.N.M.I.B.E. (2021). Quantification of soft tissue parameters from spatially resolved diffuse reflectance finite element models. *Int. J. Numer Method BioMed. Eng.* 38, e3546. doi: 10.1002/cnm.3546

Wang, X. (2020). P3 equation steady-state model of light transport in semi-infinite thick rectangular medium. *Optik* 204, 164138. doi: 10.1016/j.ijleo.2019.164138

Wang, X. (2022). P3 approximation equation of light transport in a slab medium: steady-state and time domains. *Wave. Random Complex*, 1–19. doi: 10.1080/17455030.2021.2014598

Wang, Y., Coney, C., Mcatee, C., Mccullough, G., and Goguet, A. (2022). Development of a spatially resolved technique for the measurement of effective diffusions and its application to the modelling of washcoated catalytic monoliths. *Appl. Catalysis A: Gen.* 638, 118608. doi: 10.1016/j.apcata.2022.118608

Wang, J., Fan, L., Wang, H., Zhao, P., Li, H., Wang, Z., et al. (2017). Determination of the moisture content of fresh meat using visible and near-infrared spatially resolved reflectance spectroscopy. *Biosyst. Eng.* 162, 40–56. doi: 10.1016/j.biosystemseng.2017.07.004

Wang, L., and Jacques, S. (1992). Monte Carlo modeling of light transport in multi-layered tissues in standard C. (Houston, USA: The University of Texas, M. D. Anderson Cancer Center), 1–173.

Wang, L., Jacques, S. L., Zheng, L. J. C. M., and Biomedicine, P. I. (1995). MCML—Monte Carlo modeling of light transport in multi-layered tissues. *Comput. Methods Programs Biomed.* 47, 131–146. doi: 10.1016/0169-2607(95)01640-F

Wang, A., Lu, R., and Xie, L. (2016). Finite element modeling of light propagation in turbid media under illumination of a continuous-wave beam. *Appl. Opt.* 55, 95–103. doi: 10.1364/ao.55.000095

Wang, A., Lu, R., and Xie, L. (2017). Improved algorithm for estimating the optical properties of food products using spatially-resolved diffuse reflectance. *J. Food Eng.* 212, 1–11. doi: 10.1016/j.jfoodeng.2017.05.005

Watte, R., Aernouts, B., Van Beers, R., Postelmans, A., and Saeys, W. (2016). Computational optimization of the configuration of a spatially resolved spectroscopy sensor for milk analysis. *Anal. Chim. Acta* 917, 53–63. doi: 10.1016/j.aca.2016.02.041

Wen, X., Wang, Z., and Huang, L. (2010). Measurement of myoglobin in pork meat by using steady spatially-resolved spectroscopy. *J. Agric. Eng.* 26, 375–379. doi: 10.3969/j.issn.1002-6819.2010.z2.071

Xia, J. J., Berg, E. P., Lee, J. W., and Yao, G. (2007). Characterizing beef muscles with optical scattering and absorption coefficients in VIS-NIR region. *Meat Sci.* 75, 78–83. doi: 10.1016/j.meatsci.2006.07.002

Xia, J., Weaver, A., Gerrard, D. E., and Yao, G. (2008). Distribution of optical scattering properties in four beef muscles. *Sens. Instrument. Food Qual. Saf.* 2, 75–81. doi: 10.1007/s11694-008-9032-6

Xie, D., and Guo, W. (2020). Measurement and calculation methods on absorption and scattering properties of turbid food in Vis/NIR range. *Food Bioproc. Technol.* 13, 229–244. doi: 10.1007/s11947-020-02402-3

Yang, W., Jin, X., and Gao, X. (2021). Vector radiative transfer equation for arbitrary shape particles derived from Maxwell's electromagnetic theory. *J. Quant. Spectrosc. Radiat. Transf.* 265, 107307. doi: 10.1016/j.jqsrt.2020.107307

Ye, X. J., Doi, T., Arakawa, O., and Zhang, S. H. (2021). A novel spatially resolved interactance spectroscopy system to estimate degree of red coloration in red-fleshed apple. *Sci. Rep.* 11, 21982. doi: 10.1038/s41598-021-01468-z

Yuan, H., Liu, C., Wang, H., Wang, L., and Dai, L. (2022). Early pregnancy diagnosis of rabbits: A non-invasive approach using Vis-NIR spatially resolved spectroscopy. *Spectrochim. Acta Part A: Mol. Biomol. Spectrosc.* 264, 120251. doi: 10.1016/j.saa.2021.120251

Zaytsev, S. M., Amouroux, M., Khairallah, G., Bashkatov, A. N., Tuchin, V. V., Blondel, W., et al. (2022). Impact of optical clearing on ex vivo human skin optical properties characterized by spatially resolved multimodal spectroscopy. *J. Biophotonics* 15, e202100202. doi: 10.1002/jbio.202100202

Zhang, H., Hou, Q., Luo, B., Tu, K., Zhao, C., and Sun, Q. (2022). Detection of seed purity of hybrid wheat using reflectance and transmittance hyperspectral imaging technology. *Front. Plant Sci.* 13. doi: 10.3389/fpls.2022.1015891

Zhang, L., Tian, H., Wang, L., Li, H., and Pu, Z. (2023). Selection and validation of the best detection location for hemoglobin determination by spatially resolved diffuse transmission spectroscopy. *Infrared Phys. Technol.* 133, 104839. doi: 10.1016/j.infrared.2023.104839

Zhang, L., Wang, Y., Bian, H., Wang, L., and Li, H. (2021). Optimal wavelengths selection from all points for blood species identification based on spatially resolved near-infrared diffuse transmission spectroscopy. *Infrared Phys. Technol.* 117, 103865. doi: 10.1016/j.infrared.2021.103865

Zhang, G. W., Wen, X., Wang, Z. Y., Zhao, D. J., and Huang, L. (2010). Measurement of pork tenderness by using steady spatially-resolved spectroscopy. *Spectrosc. Spectral Anal.* 30, 2793–2796. doi: 10.3964/j.issn.1000-0593(2010)10-2793-04

Zhao, W., Zhao, X., Luo, B., Bai, W., Kang, K., Hou, P., et al. (2023). Identification of wheat seed endosperm texture using hyperspectral imaging combined with an ensemble learning model. *J. Food Compos. Anal.* 121, 105398. doi: 10.1016/j.jfca.2023.105398

Zhou, Y., Fu, X., Ying, Y., and Fang, Z. (2015). An integrated fiber-optic probe combined with support vector regression for fast estimation of optical properties of turbid media. *Anal. Chim. Acta* 880, 122–129. doi: 10.1016/j.aca.2015.04.048

Zhu, Q., Guan, J., Huang, M., Lu, R., and Mendoza, F. (2016). Predicting bruise susceptibility of 'Golden Delicious' apples using hyperspectral scattering technique. *Postharvest Biol. Technol.* 114, 86–94. doi: 10.1016/j.postharvbio.2015.12.007

Zhu, Q., He, C., Lu, R., Mendoza, F., and Cen, H. (2015). Ripeness evaluation of 'Sun Bright' tomato using optical absorption and scattering properties. *Postharvest Biol. Technol.* 103, 27–34. doi: 10.1016/j.postharvbio.2015.02.007

Zude, M., Pflanz, M., Spinelli, L., Dosche, C., and Torricelli, A. (2011). Non-destructive analysis of anthocyanins in cherries by means of Lambert–Beer and multivariate regression based on spectroscopy and scatter correction using time-resolved analysis. *J. Food Eng.* 103, 68–75. doi: 10.1016/j.jfoodeng.2010.09.021

frontiers | Frontiers in Plant Science

TYPE Original Research
PUBLISHED 10 January 2024
DOI 10.3389/fpls.2023.1296473

OPEN ACCESS

EDITED BY
Qiang Lyu,
Southwest University, China

REVIEWED BY
Dimitrios Fanourakis,
Technological Educational Institute of Crete, Greece
Chu Zhang,
Huzhou University, China

*CORRESPONDENCE
Byung-Chun In
bcin@anu.ac.kr

RECEIVED 18 September 2023
ACCEPTED 19 December 2023
PUBLISHED 10 January 2024

CITATION
Kim Y-T, Ha STT and In B-C (2024) Development of a longevity prediction model for cut roses using hyperspectral imaging and a convolutional neural network.
Front. Plant Sci. 14:1296473.
doi: 10.3389/fpls.2023.1296473

Development of a longevity prediction model for cut roses using hyperspectral imaging and a convolutional neural network

Yong-Tae Kim, Suong Tuyet Thi Ha and Byung-Chun In*

Department of Smart Horticultural Science, Andong National University, Andong, Republic of Korea

Introduction: Hyperspectral imaging (HSI) and deep learning techniques have been widely applied to predict postharvest quality and shelf life in multiple horticultural crops such as vegetables, mushrooms, and fruits; however, few studies show the application of these techniques to evaluate the quality issues of cut flowers. Therefore, in this study, we developed a non-contact and rapid detection technique for the emergence of gray mold disease (GMD) and the potential longevity of cut roses using deep learning techniques based on HSI data.

Methods: Cut flowers of two rose cultivars ('All For Love' and 'White Beauty') underwent either dry transport (thus impaired cut flower hydration), ethylene exposure, or *Botrytis cinerea* inoculation, in order to identify the characteristic light wavelengths that are closely correlated with plant physiological states based on HSI. The flower bud of cut roses was selected for HSI measurement and the development of a vase life prediction model utilizing YOLOv5.

Results and discussion: The HSI results revealed that spectral reflectance between 470 to 680 nm was strongly correlated with gray mold disease (GMD), whereas those between 700 to 900 nm were strongly correlated with flower wilting or vase life. To develop a YOLOv5 prediction model that can be used to anticipate flower longevity, the vase life of cut roses was classed into two categories as over 5 d (+5D) and under 5 d (-5D), based on scoring a grading standard on the flower quality. A total of 3000 images from HSI were forwarded to the YOLOv5 model for training and prediction of GMD and vase life of cut flowers. Validation of the prediction model using independent data confirmed its high predictive accuracy in evaluating the vase life of both 'All For Love' (r^2 = 0.86) and 'White Beauty' (r^2 = 0.83) cut flowers. The YOLOv5 model also accurately detected and classified GMD in the cut rose flowers based on the image data. Our results demonstrate that the combination of HSI and deep learning is a reliable method for detecting early GMD infection and evaluating the longevity of cut roses.

KEYWORDS

cut roses, deep learning, gray mold disease, hyperspectral imaging, prediction, vase life

1 Introduction

Recently consumer interest and use of floricultural products have been growing, especially through online markets, resulting from an increase in flower sale for home use in the COVID-19 era (Bulgari et al., 2021; Gabellini and Scaramuzzi, 2022). As a large portion of floricultural plants is utilized as cut flowers, long postharvest longevity is the primary quality by which flower sales can be promoted (Vehniwal and Abbey, 2019). Although cut flower longevity cannot be readily assessed, estimates of shorter vase life commonly reduce the value of cut flowers that are shipped to international markets. The vase life of cut flowers is determined by morphological and physiological attributes, which are shaped by the interaction of preharvest conditions and genetic traits (Fanourakis et al., 2013; In and Lim, 2018). Although rose is not an ethylene-sensitive species, in some cultivars adverse effects of ethylene exposure have been reported (Macnish et al., 2010; In et al., 2017). Ethylene is a plant hormone that regulates various physiological processes, including fruit ripening and flower senescence (Wang et al., 2002). Ethylene is also produced as a product of certain agricultural commodities and industrial activities. Cut rose flowers can be exposed to increased ethylene concentrations in various situations such as storage or transport with ethylene-producing agricultural commodities (fruits or flowers that naturally produce ethylene), storage in or near industrial areas where there is high emission of the ethylene-producing substances, and the improper ventilation of the storage or transport facilities (Cape, 2003; Chang and Bleecker, 2004; Martínez-Romero et al., 2009). Dry transport is the main method employed commercially for trade because of a reduction in space (thus cheaper) and in flower bud opening (thus maturity stage is little affected) (Macnish et al., 2009). However, cut flower hydration during dry transport is reduced owing to transpiration losses, which are not compensated. In addition, some environments such as high humidity and wet conditions are characterized by increased *B. cinerea* spore density (Williamson et al., 2007; Friedman et al., 2010). This increased spore density is not apparent at harvest, but later on, problems appear.

Therefore, the vase life of cut roses commonly ends during the early stages of flowers' development, and reliably predicting their lifespan has not been possible. Consequently, consumers are dissatisfied and flower utilization is reduced (Reid et al., 1996; Vehniwal and Abbey, 2019). Thus, the development of longevity prediction techniques is a high priority to assure the ornamental period of cut flowers for the customers, as this can be incorporated into the existing system for quality grading of cut flowers. Moreover, the vase life prediction system can improve efficiency in flower supply chains as well as provide consumers with relevant flower products according to their use. For instance, different batches of cut flowers can be sorted based on the vase life potential in the packaging house. The flower batches with short vase life are priced lower and traded shortly, whereas the batches with long vase life are priced higher in the markets and can be stored for longer time before the distribution. Furthermore, the vase life prediction model in cut flowers offers benefits ranging from quality assurance and supply chain optimization to environmental sustainability and economic efficiency. It aligns with the boarder goals of the floral industry, aiming to deliver high quality products while minimizing waste and environmental impact.

Few attempts had so far been made to devise effective methods to predict and guarantee postharvest longevity of cut flowers. Staby and Cunningham (1980) reported a method to estimate the vase life of cut carnation based on the ethylene level using gas chromatography. However, vase life prediction using this method is not suitable in ethylene-insensitive flowers and might be less accurate in the early stage of postharvest. Tromp et al. (2012) developed a method to predict the remaining vase life of cut roses using the degree-days model during storage and transportation at a constant. However, this method may be of limited use if the biological variance is high or the temperature of storage and transportation is outside the optimum range (2-6 °C).

We developed previously artificial neural network models to predict and assure the vase life of three rose cultivars based on thermal image analysis. Although the prediction accuracy of the models was quite high, the application of this method was limited because the cut roses used for the prediction model did not undergo various postharvest conditions that influence the vase life of cut flowers, such as dry transport, exposure to ethylene, or high density of mold spore during storage and transport (In et al., 2009; In et al., 2016a). Thus, to enhance the model performance for practical application in the vase life guarantee, it is further necessary to detect plant status rapidly and to use extensive data processing for complex data, such as artificial intelligence or machine learning.

Recently, a non-destructive method such as hyperspectral imaging (HSI) has been widely used to evaluate various factors related to plant physiology and stress conditions in multiple horticultural crops (Behmann et al., 2014; Liu et al., 2015; Lowe et al., 2017; Veys et al., 2019; Ramamoorthy et al., 2022; Wieme et al., 2022). HSI uses a hyperspectral camera to capture images of plants in a wide range of light wavelengths (Lowe et al., 2017; Lay et al., 2023). By analyzing the reflectance of horticultural products in different wavelengths, HSI can extract detailed information about the morphological and physiological properties of plants, including disease infection, nutritional deficiencies, ripeness, and defects of fruits and vegetables, etc (Liu et al., 2015; Wieme et al., 2022). The development of spectral imaging techniques has required suitable regression models to analyze spectral data. Machine learning techniques based on algorithms have been applied to construct classification and regression models for HSI to predict and evaluate the quality of vegetables and fruits (Zhang et al., 2016; Rahman et al., 2017; Ji et al., 2019). However, the machine learning algorithms only performed a screening process on the spectral bands (Zhang et al., 2016). In recent years, deep learning, a subset of machine learning, has been widely used in agriculture, industry, and medics because it can learn features automatically from a large dataset of images (Guo et al., 2016; Tian et al., 2020). This technique was used in building hyperspectral imaging correction models for prediction and classification. Convolutional neural networks (CNNs), a type of deep learning algorithm, can rapidly and accurately classify the quality of agricultural products and identify potential factors affecting their appearance or shelf life without being influenced by personal biases or subjective opinions (LeCun

et al., 2015; Kamilaris and Prenafeta-Boldú, 2018; Cravero et al., 2022). In the last decade, CNNs have been increasing employed in plant phenotyping community. They have been very effective in modeling complicated concepts, owing to their ability of distinguishing patterns and extracting regularities from data (Nasiri et al., 2021; Taheri-Garavand et al., 2021). You Only Look Once version 5 (YOLOv5), a type of CNN, is a state-of-the-art deep learning algorithm that was used to classify agricultural products with high accuracy even when source images are poor quality or contain multiple features (Yao et al., 2021; Ahmad et al., 2022). To classify agricultural products by using YOLOv5, the algorithm must first be trained on a large dataset of labeled images (Redmon et al., 2016; Yao et al., 2021). YOLOv5 can also perform real-time classification, which is important for rapidly classifying large quantities of horticultural products (Zhang et al., 2021; Li et al., 2022). HSI and deep learning techniques have been widely applied to predict postharvest quality and shelf life in multiple horticultural crops such as vegetables, fruits, and mushrooms (Taghizadeh et al., 2011; Mo et al., 2015; Susič et al., 2018; Sun J et al., 2021; Wieme et al., 2022; Xiang et al., 2022); however, there are few studies showed the application of these techniques to evaluate the quality issues of cut flowers (Stead et al., 2018; Sun X et al., 2021). Therefore, this study aimed to develop a rapid and effective method to predict the longevity of cut roses based on HSI and deep learning algorithms. To identify light wavelengths that are closely correlated with plant physiological states (GMD and petal wilting) using HSI, cut flowers underwent either water stress, ethylene exposure, or *B. cinerea* inoculation before storage. YOLOv5 was adopted for processing the extensive image data by HSI in order to develop vase life prediction models for cut flowers. In the present study, the flower bud of cut roses was chosen for HSI measurement and the development of the vase life prediction model. This selection allows for imaging from the top of entire batches of cut flowers. Furthermore, the results obtained in this study are not confined solely to hydration status; they also contribute to the vase life prediction for cut rose flowers.

2 Materials and methods

2.1 Plant materials

Cut roses 'All For Love' and 'White Beauty'(*Rosa hybrida* L.) were cultivated and harvested in a commercial greenhouse in Guksong, Jeollanam-do, South Korea. Rose plants were dripirrigated with a liquid nutrient solution containing NH_4NO_3 (44.93 $_g$ L^{-1}), $Ca(NO_3)_2$ $4H_2O$ (17.47 g L^{-1}), KNO_3 (1.63 $_g$ L^{-1}), KH_2PO_4 (12.04 $_g$ L^{-1}), $MgSO_4$ $7H_2O$ (27.04 g L^{-1}), and a small volume of other trace substances. The symptomless rose flowers were collected and harvested at the commercial stage (outer petals bent out) (Harkema et al., 2013). After harvest, cut flowers were either wet transported (WT) in tap water or dry transported (DT) without the water to the laboratory within 4 h. At the laboratory, all cut roses were placed in a controlled environment room at 23 ± 1 °C and at a relative humidity (RH) of 50 ± 2% for HSI analysis. After the HSI, the cut flowers were exposed to ethylene or inoculated with *B. cinerea* and followed by storage at 10 ± 1 °C and RH of 50 ± 5% under dark conditions for 3 d for transport treatments (In et al., 2016b).

2.2 Ethylene exposure

Cut flowers in WT were held in distilled water and those in DT were placed in buckets without water and enclosed in the treatment chamber (462 L) at 23 ± 1 °C under dark conditions. Distilled water was used, though less common from practical stand point, since the tap water composition largely depends on the season, and the location (Amadi-Majd et al., 2021). Ethylene (10%) was injected into the chamber to achieve a final concentration of 2 μL L^{-1}. Three beakers containing 200 mL of 1M NaOH were placed in the treatment chamber to neutralize CO_2 released by the flower respiration during the ethylene treatment. After every 12 h of ethylene exposure, the treatment chamber was opened for 2–3 h for HSI and then closed and re-injected with 2 μL L^{-1} ethylene. Three days after the transport treatments, cut flowers were taken out from the chamber for vase life evaluation and HSI.

2.3 *B. cinerea* inoculation in cut roses

B. cinerea (KACC40573) was isolated from infected rose flowers in the Korean Agricultural Culture Collection (KACC), National Institute of Agricultural Sciences. For a pure culture, *B. cinerea* conidia were grown in potato dextrose agar (PDA, Difo Laboratories, Detroit, MI, USA) at 25 ± 1 °C for 14 days. *B. cinerea* conidial suspension was obtained by dropping 10 mL of distilled water into a culture petri dish and then gently sweeping the fungal colony surface with a sterile loop. The conidial clumps were removed from the obtained suspension by gently filtering with sterile gauze. Afterward, the concentration of conidia suspension was adjusted to 10^5 conidia mL^{-1} with sterile water for the experiment.

WT and DT flowers were inoculated by spraying with 30 mL of the conidial suspension (10^5 conidia mL^{-1}). Non-inoculated cut roses were sprayed with sterile water (30 mL). After inoculation of *B. cinerea*, the rose flowers were then placed in the storage chamber (at temperature 10 ± 1 °C and RH of 50 ± 5%) under dark conditions for 3 d to simulate export conditions. After the transport treatments, cut flowers were set up for vase life and disease progression evaluation and HSI.

2.4 Evaluation of vase life and gray mold disease

After three days of the export simulation, twenty-five cut roses in each treatment were trimmed to a length of 45 cm with three upper leaves. Each cut flower was placed into a glass jar containing distilled water (450 mL) and maintained at the temperature (23 ± 1 °C), RH of 50 ± 2%, and light intensity at 20 μmol m^{-2} s^{-1} (a photoperiod of 12 h) supplied by fluorescence tubes for GMD progression and vase life assessment.

Changes in the postharvest quality of cut roses were determined by measuring relative fresh weight and water uptake daily at 10:30 am. Water balance (WB) of cut flowers was calculated from changes in fresh weight, water uptake, and daily transpiration. The vase life of cut roses was evaluated daily by the assessment criteria for *Rosa* (VBN, 2014). Cut roses were considered to have reached the end of their postharvest life when flowers showed at least one or more of the following senescence symptoms: pedicel bending (neck angle greater than 45°), petal drying (≥ 50% of petals show dryness); wilting of petal and leaf (≥ 50% of petals or leaves loss their turgor), petal abscission (a drop of three or more petals), leaf abscission and yellowing (≥ 50% leaf drop and yellowing), bluing (≥ 50% blue petals) (Fanourakis et al., 2015; Fanourakis et al., 2016). In addition, the vase life of cut roses was considered to end when cut flowers showed severe GMD symptoms in the petals. The progression of GMD by *B. cinerea* was evaluated based on the disease index as described in the previous study (Ha et al., 2022).

2.5 Fungal biomass and gene expression analysis

Fungal genomic DNA (gDNA) was extracted from the gray mold mycelia collected from infected petals by using i-genomic BYF DNA Extraction Mini Kit (INTRON Biotechnology Inc., Gyeonggi-do, South Korea). Total RNA was isolated from 200 mg of rose petals by using the GeneJET plant RNA purification Mini Kit (Thermo Fisher Scientific Baltics, Vilnius, Lithuania). cDNA was synthesized from 1 μg of total RNA using XENO-cDNA Synthesis Kit (CELL TO BIO, Gyeonggi-do, South Korea) and performed in a Bio-Rad PTC-100 Programmable Thermal Controller (MJ Research Inc., Hercules, CA, USA) as per the instruction manual. Then, fungal biomass (evaluated by *Bc3* from gDNA) and the transcript levels of the ethylene biosynthesis gene (*RhACO1*), aquaporin-related gene (*RhTIP1*), and senescence-induced gene (*RhSIG*) in petals of cut roses were analyzed using the BIO-RAD CFX Connect Real-Time System (Life Science, Hercules, CA, USA). *B. cinerea* actin A (*BcactA*) and *Rosa hybrida* actin 1 (*RhACT1*) genes were used as an internal control. The primer sequences used for quantitative real-time PCR (qRT-PCR) are listed in Table 1. The qRT-PCR reaction setting and conditions for gene expression analyses have been indicated previously (Ha et al., 2022).

2.6 Hyperspectral image acquisition

The visible/near-infrared (VIS-NIR) hyperspectral camera system was set with an IMEC SNAPSCAN camera (3650x2048 pixel) (IMEC, Leuven, Belgium, www.imec-int.com) with 150 spectral bands and a spectral range of 470–900 nm. This system was connected to a computer (Intel (R) Core (TM) I7-1165G7 CPU @ 2.8 GHz). Images of cut roses were acquired using the HSI in reflection mode and were constructed under a controlled environment room (23 ± 1 °C and RH of 50 ± 2%). The VNIR light source was provided by 4 halogen Osram lamps with 20W HT spot and color temperature of 2800 K (OSRAM, Munich, Germany). The halogen lamps provide 350-2500 nm light with a power of 20 W. The distance between the cut rose flowers and the lens was set to 50 cm, and the angle between the lamps and camera was set at 45° to provide enough light to the imaging area for image acquisition. The exposure time of the hyperspectral camera shooting was set to 2 milliseconds. The halogen lamps were run for 15 min to reach a stable state temperature and then a 95% reflection standard was calibrated before conducting reflection measurements of the cut roses. Data acquisition and extraction were performed using the IMEC HSI Snapscan software version 1.8.1.1 (IMEC, Leuven, Belgium).

2.7 Image processing model

A dataset of images of cut roses was used to process disease detection and vase life prediction by using deep learning system YOLOv5 version 6.2 (GitHub, San Francisco, USA). The dataset consisted of 3000 images collected from the hyperspectral system, with 1500 disease-infected cut roses and 1500 non-disease-infected cut roses. The images were resized to 640x640 pixels and the disease-infected areas in the images were annotated with bounding boxes using MAKE SENSE (Figures 1A–C). The annotation process was done by a trained 1 annotator is familiar with disease-infected cut roses to ensure consistency and accuracy. The YOLOv5 architecture implemented in Python using the PyTorch library was used for object detection. The YOLOv5x model was implemented using the GitHub library and was trained on a computer with a CUDA-enabled GeForce RTX 3080 graphics card for 50 epochs. To evaluate the performance of gray mold disease detection in cut roses, metrics including precision (P), recall (R), mean average precision (mAP), and F1-score (F1) were

TABLE 1 List of genes and primers used for qRT-PCR analysis in this study.

Gene (accession number/reference)	Forward primer	Reverse primer
Bc3 (Suarez et al., 2005)	5'-GCTGTAATTTCAATGTGCAGAATCC-3'	5'-GGAGCAACAATTAATCGCATTTC-3'
BcactA (Chagué et al., 2006)	5′-CCCAATCAACCCAAAGCTCAACAG-3′	5′-CCACCGCTCTCAAGACCCAAGA-3′
RhACO1 (AF441282.1)	5′CGTTCTACAACCCAGGCAAT-3′	5′-TTGAGGCCTGCATAGAGCTT-3′
RhTIP1 (KF985188.1)	5'-TCTCTCCTACGTGGCATCCT-3'	5'-GACCACCTCTGCTTTTGCTC-3'
RhSIG (S80863.1)	5'-CCGACACAAGAACCTTGGAT-3'	5'-TCTTCCGTGTACACCACCAA-3'
RhACT1 (KC514918.1)	5′-GTTCCCAGGAATCGCTGATA-3′	5′-ATCCTCCGATCCAAACACTG-3′

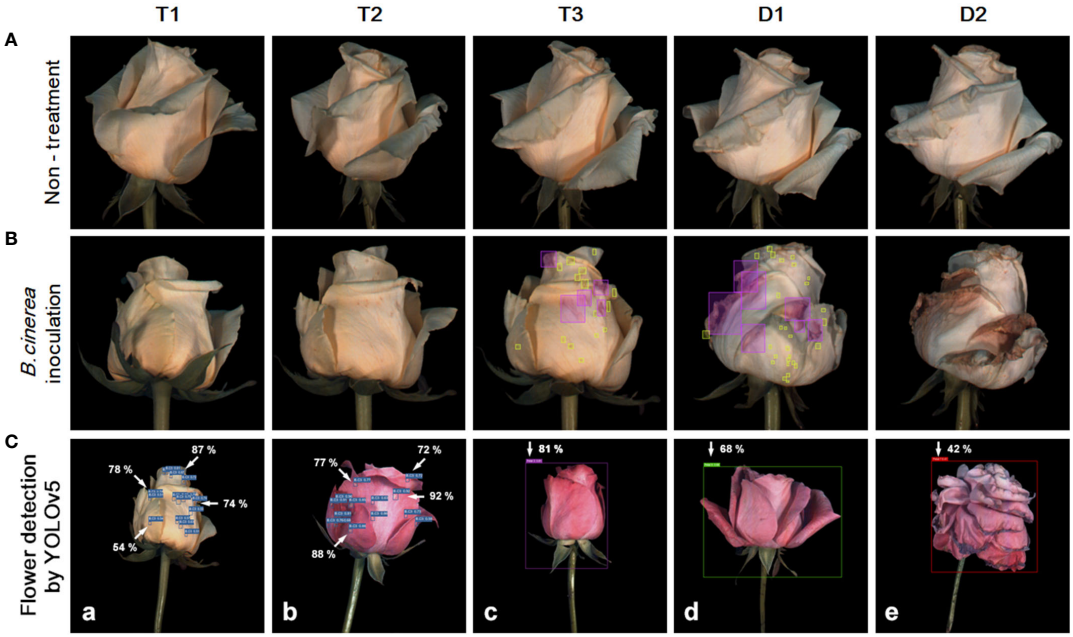


FIGURE 1
Development of flower opening and gray mold disease (GMD) in 'White Beauty' cut roses during transport and vase life **(A, B)**. The cut flowers were untreated (none treatment) or sprayed with *B. cinerea* conidia suspension (inoculation) and the images of flowers were taken on days 1-3 of transport (T1-T3) and days 1-2 (D1-D2) of vase period. The green and pink boxes indicate the annotation of the GMD emergence spots by bounding boxes for deep learning analysis **(B)**. Detection of GMD and petal wilting in 'White Beauty' **(a)** and 'All For Love' **(b–e)** by YOLOv5 **(C)**. The arrows and numbers in the flower images indicate the GMD spots and the probability (%) of GMD calculated by YOLOv5 **(a, b)**. The bounding boxes in purple, green, and red generated by annotation tool MAKE SENSE indicate petal wilting and opening levels of the flowers at T0, T3, and D1. The percentage numbers in the images indicate the probability of the specific wilting and opening stages, as calculated by YOLOv5 **(c–e)**.

used in the present study. The target confidence threshold was 0.5 and the Intersection over Union (IOU) at the time of testing was 0.5. The P, R, mAP, and IOU are calculated as follows:

$$P = \frac{TP}{TP + FP}$$

$$R = \frac{TP}{TP + FN}$$

$$mAP = \frac{\sum_{i=1}^{k} AP_i}{k}$$

$$F1 = 2 \times \frac{P \times R}{P + R}$$

$$IOU = \frac{Area\ \ of\ \ Overlap}{\text{Area of Union}}$$

Where TP, FP, and FN are the numbers of true positive cases, false positive cases, and false negative cases. The specific network structure of YOLOv5x is shown in Figure 2.

To identify the most appropriate image processing model, we also evaluated the performance of two more object detection models: Faster R-CNN and Single Shot Muli-Box Detector (SSD). We utilized the cut rose image dataset, which includes 588 images across 21 categories, showcasing various senescence symptoms. The dataset was partitioned into 70% for training, 15% for validation, and 15% for testing. We tailored the input image sizes to meet the requirements of each model: 640x640 pixels for Faster R-CNN and 512x512 pixels for SSD. All models were implemented using the PyTorch open-source deep learning framework. Each model underwent training with identical hyperparameter settings, including a learning rate set to 0.001, a batch size of 16, and training for a total of 50 epochs.

To identify initial disease symptoms and wounded spots, we used an image region extraction pre-processing step using the YOLOv5 object detection algorithm. The flower objects within the images were identified and boxed with a rectangular frame. The objects in the bounding boxes were then precisely cropped and the small spots were detected from the images by the image pre-processing system as shown in Figure 3.

A random forest classification model using the object values detected by the YOLOv5x was used to predict the vase life of cut roses. We used a dataset of 200 cut roses corresponding with vase life labels ranging from 1 to 8 d. To optimize the performance of the random forest model, the object values were grouped into feature sets of 1 to 12, based on the importance ranking of the 12 features. The feature sets were constructed by iteratively adding the next most important feature to the previous set until 12 features were included. The dataset was split into training and testing sets using an 80:20 ratio, with stratified sampling to ensure that both sets have a similar distribution of the vase life labels. The random forest model was trained using the training set, with hyperparameters optimized using grid search and cross-validation. The optimized hyperparameters included 100 trees, a maximum depth of 10, and minimum samples required to split a node of 2.

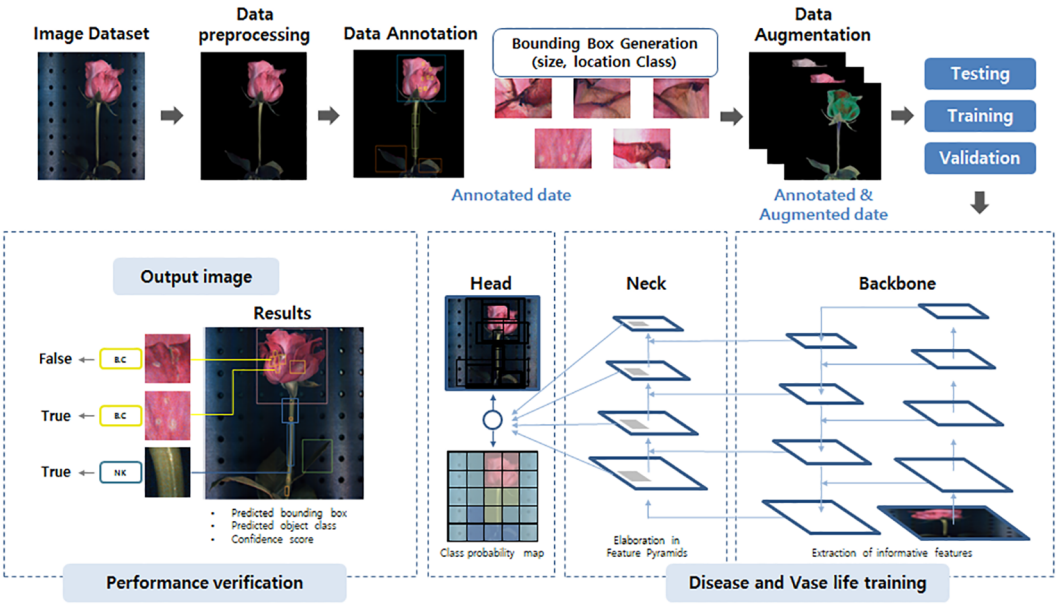


FIGURE 2
The YOLOv5 network model structure schematic used in this study. The image dataset is first pre-processed, annotated, and undergoes data augmentation to enhance its quality and quantity. The prepared image data is then forwarded to the backbone, the neck, and the head of the model for training and prediction of disease and vase life. Finally, the prediction performance of the models was estimated based on the object-detection values by YOLOv5 system.

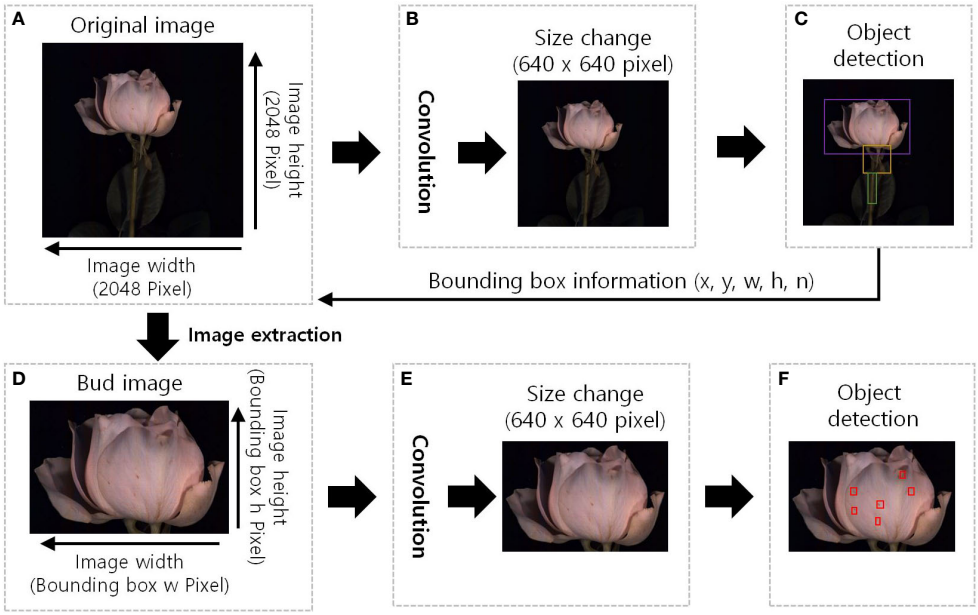


FIGURE 3
The GMD detection image region extraction pre-processing diagram of the YOLOv5. **(A)**, original HSI with a resolution of 2048 pixels in width and height; **(B)**, the original image was resized to 640x640 pixels, and convolution operations were performed to extract features. **(C)**, detect the object and the bounding box information surrounding the object (x, y, w, h, n) was extracted, x and y: position of the bounding box, w: width, h: height, and n: identification number of the object. **(D)**, bud image is the extracted image of the detected object region, image size is the width and height of the bounding box; **(E)**, resized of 640x640 pixels to standard the input size and additional feature extraction; and **(F)**, red boxes in indicate the detection of disease and wounded spots.

The output of the vase life was classed into two categories as over 5 d (> 5 d) and under 5 d (≤ 5 d) based on the total scores evaluated by gray mold disease (GMD) severity, GMD development weighted value, petal wilting level, and flower opening as shown in Table 2 and Supplementary Figure 1. The scores of quality factors used to predict the vase life of cut roses in Table 2 were calculated based on the incidence of the vase life terminated factors and GMD disease (Supplementary Figure 2). The GMD development weighted value was determined by the growth speed of the disease in petals. The GMD development speed was accelerated by *B. cinerea* inoculation and ethylene treatment and also increased in 'White Beauty' compared to 'All For Love' (Supplementary Figure 3). This evaluation was based on the previous findings showing that ethylene and water stress influenced the progression of GMD in cut roses during transport (Harkema et al., 2013; Ha et al., 2022).

2.8 Experimental design and statistical analysis

Twenty-five cut roses were used for each treatment. Experiments on vase life and disease evaluation were performed with 10 replicates (one cut flower per replicate). HSI analyses were performed with 6 cut flowers. The remaining 9 cut flowers were used for fungal biomass and gene expression analysis. qRT-PCR analysis was conducted with 3 biological replicates. Data were subjected to analysis of variance (ANOVA) or simple linear regression analysis at $p< 0.05$ using SPSS version 22.0 (IBM, Somers, NY, USA). Data are presented as the mean ± standard error (SE). The experiments were performed twice in both rose cultivars.

3 Results

3.1 Transport treatments influence vase life, water status, disease infection, and total reflectance of cut roses

WT treatment extended the vase life of cut roses compared to other treatments (Figures 4A, B). WT yielded the longest vase life in both 'All For Love' (5.3 d) and 'White Beauty' (5.2 d) varieties of cut roses (Figures 4A, B). Conversely, DT, ethylene, and *B. cinerea* treatments significantly reduced the vase life of both cultivars (Figures 4A, B). Similarly, changes in both cultivars' capacity to maintain WB mirrored the changes in vase life in response to the different transport treatments. (Figures 4C, D).

The first visual symptoms of gray mold disease (GMD) were observed on day 1 (T1) of transport in WT+E and WT+B flowers in both rose cultivars (Figures 4E, F). WT+E and WT+B treatments most increased GMD severity in the flower petals during vase periods (Figures 4E, F). Although DT reduced the vase life of cut roses, due to water stress caused by an early disruption of water balance, this transport method delayed GMD growth in the flower petals (Figures 4E, F). In particular, 'All For Love' DT flowers showed no GMD symptoms after transport treatment (Figure 4E).

Mean spectral reflectance curves of the cut roses in the wavelength range 470-900 nm obtained using the HSI on the first day (D1) of the vase period are shown in Figures 4H, G. The size and shape of flower buds did not influence the reflectance of wavelength in cut rose flowers (Supplementary Figure 4). The overall spectral patterns induced by the two treatments were similar for both cultivars. The reflectance of wavelength (RW) in WT flowers was higher than those of other flowers (Figures 4H, G), whereas that of DT, DT+E, DT+B, WT+E, and WT+B flowers was relatively low and corresponded with the decline in both vase life and capacity to maintain water balance, as well as and the increase in GMD index (Figures 4H, G). The distinct differences in RW in the 470-680 nm range (RW 470/680) in both rose cultivars perhaps show the relation of the spectrums to the susceptibility to the gray mold of the cut flowers (Figures 4H, G). Conversely, the differences in RW in the 700-900 nm range (RW 700/900) in both rose cultivars may be correlated with the flower responses to water stress and ethylene (Figures 4H, G).

3.2 Changes in spectrum curves, fungal growth, and relative expression of genes involved in ethylene biosynthesis, water stress, and senescence of cut roses

Changes in spectral reflectance of cut roses in each treatment group (solid lines) were analyzed throughout the transport and vase

TABLE 2 The scores of quality factors used for to predict the vase life of cut roses using YOLOv5.

GMD severity[x]		GMD weighted value[y]			Petal wilting[z]		Vase life[w]	
Level	Score	Treatment/cultivar		Score	Level	Score	Total score	Output
1	0		None	0	1	0	> 60	≥ 5 D
2	20		*B. cinerea*	20	2	20	< 61	< 5 D
3	40		Ethylene	20	3	40		
4	100	Culti-var	'All For Love'	0	4	100	< 100	Exclusion
			'White Beauty'	20				
Total score = 100 – (GMD severity + GMD weighted value + Petal wilting)								

[x]GMD, gray mold disease.
The severity of GMD was evaluated by the area (%) of the disease symptom in rose petals as follows: 1, none; 2, ≤ 3%; 3, 3-10%; and 4, 11-50%.
[y]The weighted value of GMD is the disease developmental speed in rose petals influenced by the treatments and cultivars.
[z]Petal wilting was influenced by water stress and flower opening. It was calculated using four levels as follows: 1, none; 2, slight wilting; 3, moderate wilting; 4, severe wilting.
[w]The total score of vase life was the sum of the scores graded by the quality factors. Vase life of cut roses was classified in two categories: over 5 d (+5D) and under 5 d (-5D) based on the total score. If the total score was ≥ 100, the cut flowers were excluded from the vase life evaluation and classified into the defective group.

10.3389/fpls.2023.1296473

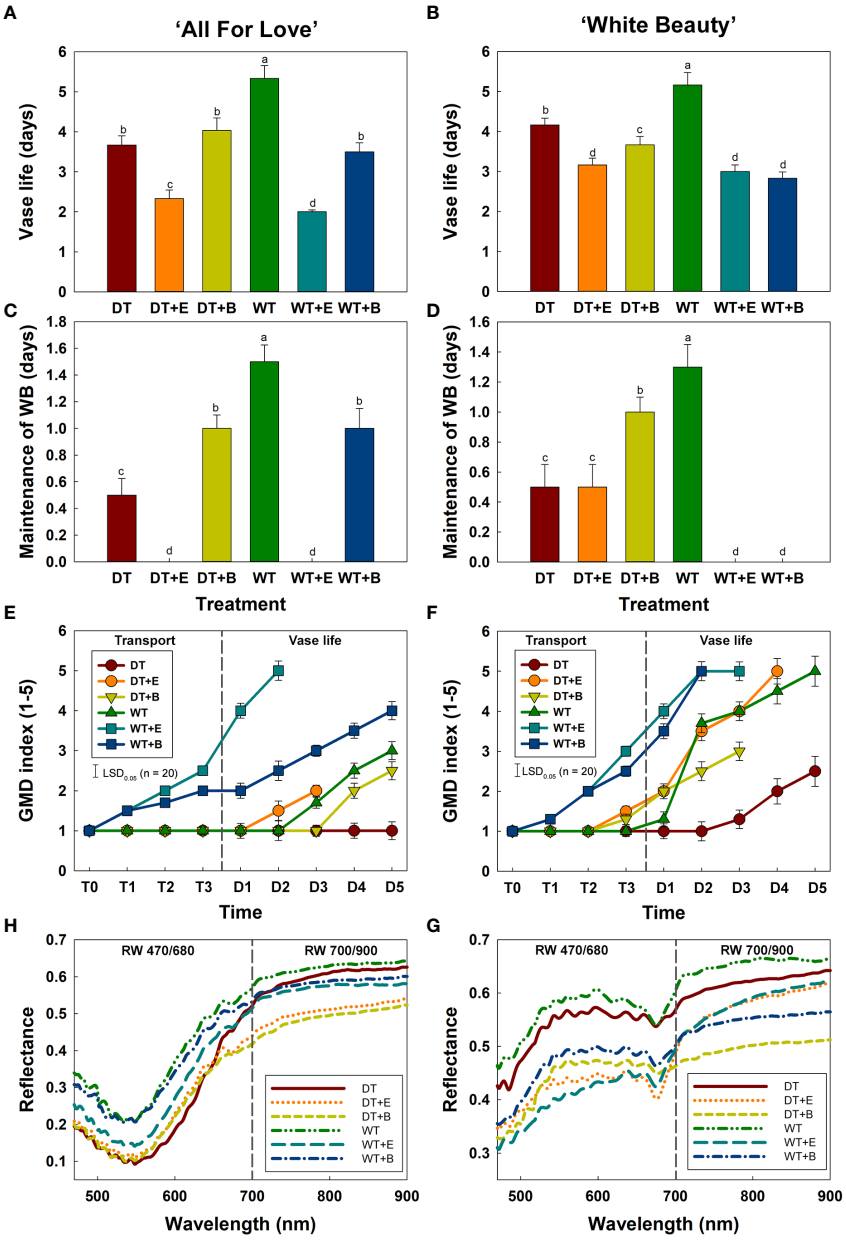


FIGURE 4
Changes in physiological characteristics of cut roses during transport and vase period. **(A, B)**, vase life; **(C, D)**, maintenance of water balance; **(E, F)**, gray mold disease (GMD) index; **(G, H)**, reflectance of wavelength (RW) in cut roses. DT, dry transport; DT+E, ethylene exposure before DT; DT+B, *B. cinerea* inoculation before DT; WT, wet transport; WT+E, ethylene exposure before WT; WT+B, *B. cinerea* inoculation before WT. GMD was evaluated on days 1-3 of transport (T1–T3) and days 1-5 (D1–D5) of vase period. GMD index was classified into five levels as 1, none; 2, slight symptoms (≤ 3%); 3, moderate symptoms (3-10%); 4, severe symptoms (11-50%); and 5, death of plants (> 50%). RW was detected in cut roses on the first day (D1) of vase period. RW 470/680 and 700/900 indicate the wavelengths from 470 to 680 nm and 700 to 900 nm, respectively. Data are shown as means ± SE (n = 20). Different letters above bars indicate statistically significant differences among treatments at $p = 0.05$ based on Duncan's multiple range test.

periods, and the corresponding changes in GMD growth (*Bc3* level) and the relative expression of genes related to ethylene biosynthesis (*RhACO1*), water stress (*RhTIP1*), and senescence induction (*RhSIG*) were also detected in the petals (bar charts) (Figure 5).

The changes in total spectral reflectance in both rose cultivars after transport treatments are shown in Supplementary Figure 5. In the various treatment groups of 'All For Love' roses, RW 470/680 during transport (T0) varied in correlation with the level of fungal biomass in petals (Figure 5A). Ethylene, *B. cinerea* inoculation, and WT induced high *Bc3* levels rapidly in cut roses while DT reduced *Bc3* levels in rose petals (Figure 5A). Thus, RW470/480 in DT+E, DT+B, WT, WT+E, and WT+B flowers rapidly decreased due to *B. cinerea* growth after transport treatments (Figure 5A). In contrast, RW 470/480 in DT roses changed only slightly during vase periods (Figure 5A). In the case of the 'White Beauty', these flowers are particularly susceptible to GMD; thus, the fungal biomass (*Bc3*

10.3389/fpls.2023.1296473

level) emerged in the petals of all cut roses early at D1 (1st day of the vase period). Consequently, the reduction in RW 470/680 was similar in all flowers (Figure 6A).

Ethylene exposure induces higher mRNA levels of the ethylene biosynthesis-related gene *RhACO1* in rose petals (In et al., 2017). Moreover, both ethylene and water stress reduced the expression levels of *RhTIP1*, an aquaporin-related gene (Xue et al., 2009; Ha et al., 2021). These changes induced early senescence symptoms in cut roses by stimulating the expression of senescence-induced genes (Figures 5B–D, 6B–D). In all flowers, a decrease in RW 700/900 corresponded to increased mRNA levels of *RhACO1* and *RhSIG* (a senescence-induced gene) and a decrease in *RhTIP1* expression in petals (Figures 5B–D, 6B–D). At the later stage of the vase period (D4), the death of petal tissues due to GMD or senescence caused a decline in the spectral reflectance of all cut flowers (Figures 5A–D, 6A–D).

To confirm the above results, we extracted the RW 470/680 and RW 700/900 from petals based on GMD index differences (Figures 7A, C) and petal wilting level due to water stress or ethylene exposure (Figures 7B, D). Subsequently analysis,

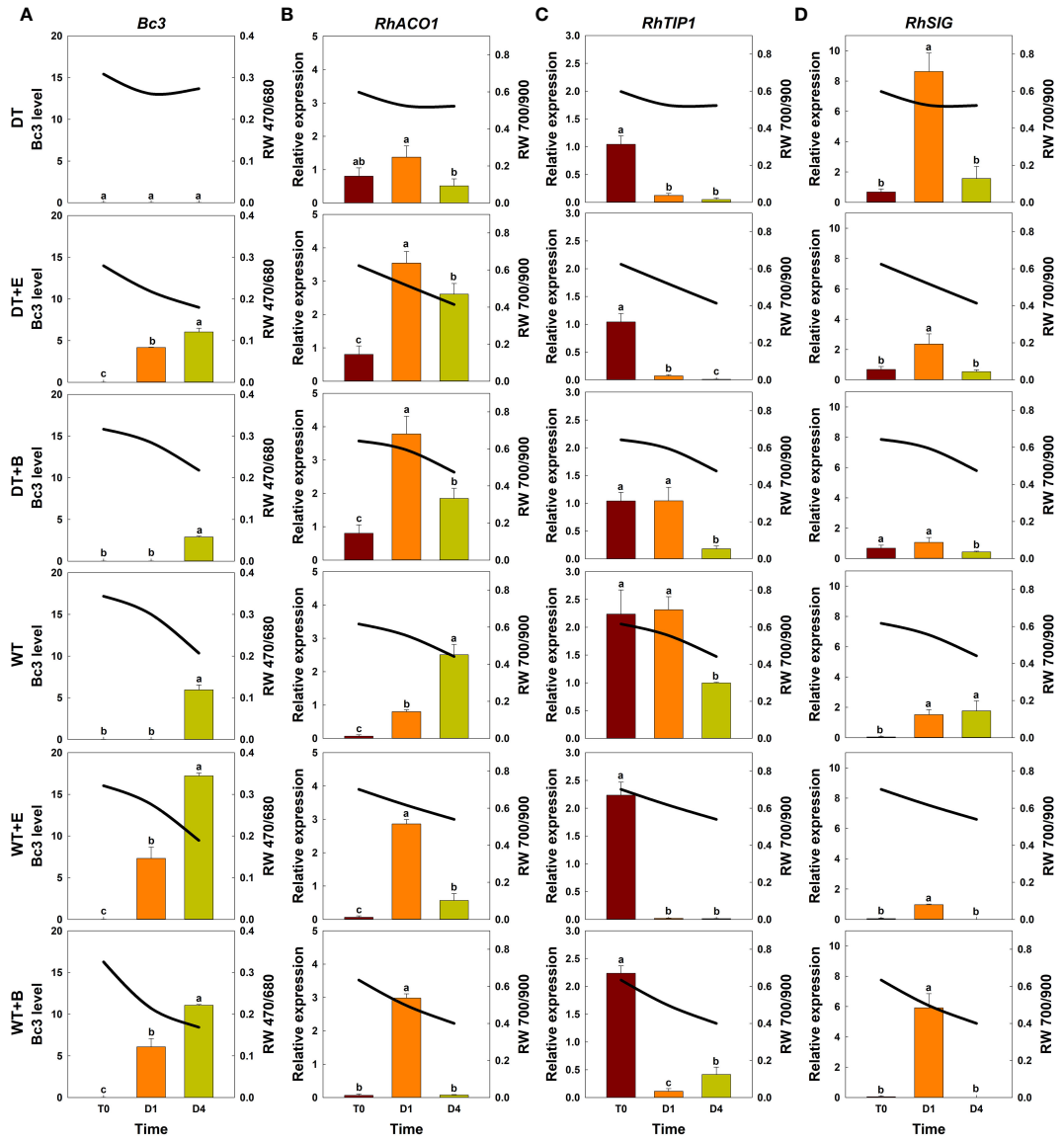


FIGURE 5
Effect of transport treatments on changes in *B. cinerea* infection level, relative expression of genes related to flower longevity, and average reflectance of wavelength (RW) in 'All For Love' cut roses. RW 470/680 and 700/900 indicate the wavelengths of 470 to 680 nm and 700 to 900 nm, respectively. *Bc3*, fungal biomass in rose petals **(A)**; *RhACO1*, ethylene biosynthesis gene **(B)**; *RhTIP*, aquaporin-related gene **(C)**; and *RhSIG*, senescence-induced gene **(D)**. RW, fungal biomass, and gene expression level in cut roses were analyzed on day 0 of transport (T0) and on days 1 (D1) and 4 (D4) of the vase period. DT, dry transport; DT+E, ethylene exposure before DT; DT+B, *B. cinerea* inoculation before DT; WT, wet transport; WT+E, ethylene exposure before WT; WT+B, *B. cinerea* inoculation before WT. The solid line represents the average reflectance of wavelength. The bar charts represent the *Bc3* level, and relative expression of genes related to flower longevity. Data are shown as means ± SE (n = 20 for RW data, 6 for gene expression data). Different letters above bars indicate statistically significant differences among treatments at $p = 0.05$ based on Duncan's multiple range test.

10.3389/fpls.2023.1296473

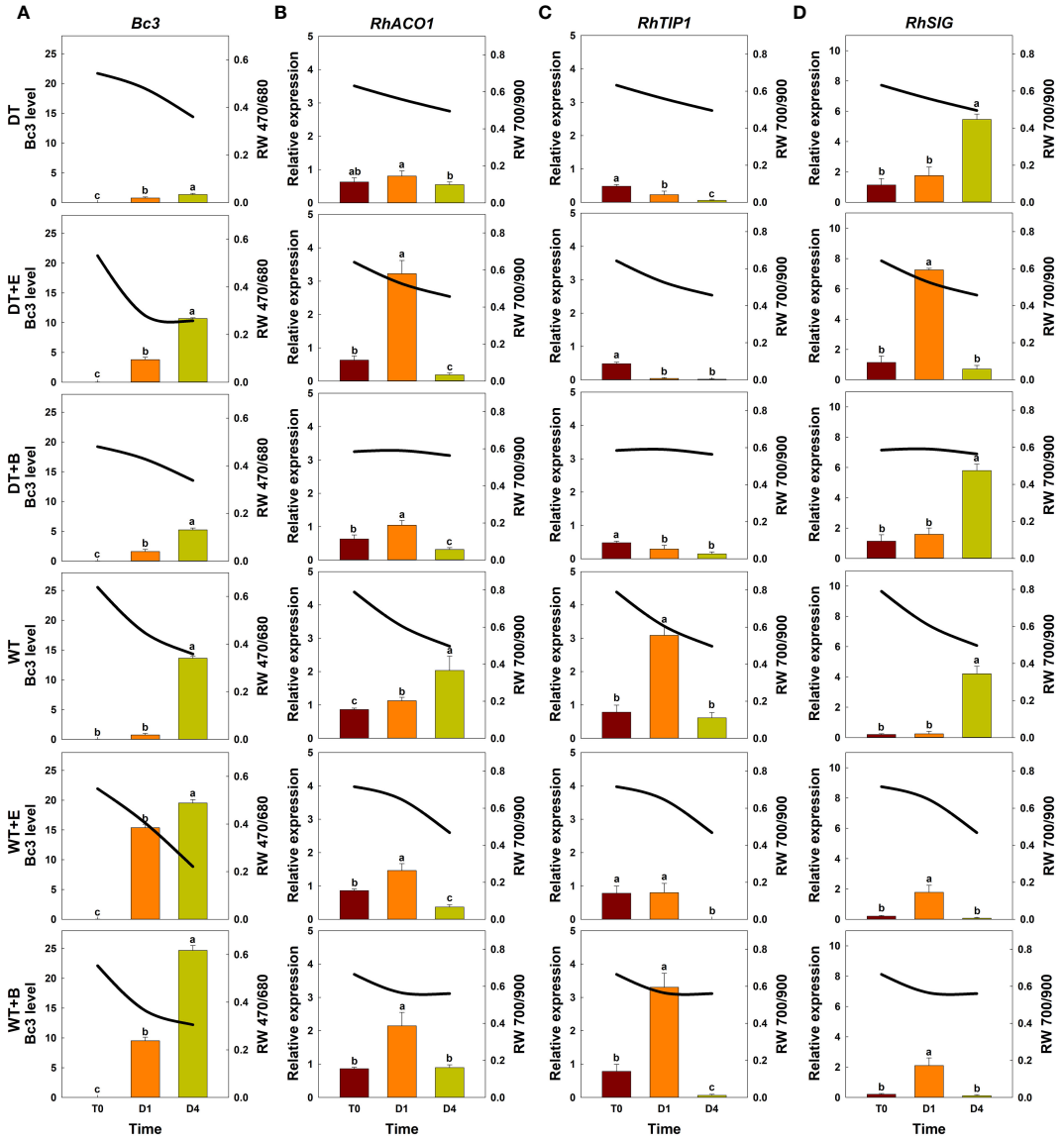


FIGURE 6
Effect of transport treatments on changes in *B. cinerea* infection level, relative expression of genes related to flower longevity, and average reflectance of wavelength (RW) in 'White Beauty' cut roses. RW 470/680 and 700/900 indicate the wavelengths of 470 to 680 nm and 700 to 900 nm, respectively. *Bc3*, fungal biomass in rose petals **(A)**; *RhACO1*, ethylene biosynthesis gene **(B)**; *RhTIP1*, aquaporin-related gene **(C)**; and *RhSIG*, senescence-induced gene **(D)**. RW, fungal biomass, and gene expression level in cut roses were analyzed at day 0 of transport (T0) and at days 1 (D1) and 4 (D4) of the vase period. DT, dry transport; DT+E, ethylene exposure before DT; DT+B, *B. cinerea* inoculation before DT; WT, wet transport; WT+E, ethylene exposure before WT; WT+B, *B. cinerea* inoculation before WT. The solid line represents the average reflectance of wavelength. The bar charts represent the *Bc3* level, and relative expression of genes related to flower longevity. Data are shown as means ± SE (n = 20 for RW data, 6 for gene expression data). Different letters above bars indicate statistically significant differences among treatments at $p = 0.05$ based on Duncan's multiple range test.

employing a one-way ANOVA test for each RW, identified RW 600-680 nm in 'All For Love' and at RW 500-650 nm in 'White Beauty', with notably high *p*-values, closely related to GMD symptom severity (Figures 7A, C). Additionally, high *p*-values at RW 700-900 nm indicated distinctions in petal wilting (Figures 7B, D). Whereas, *p*-values were low at RW 700/900 and RW 470/680, which are related to GMD severity (Figures 7A, C) and petal wilting levels (Figures 7B, D). These results indicate that RW 470/680 and RW 700/900 are closely correlated to GMD and other stress conditions respectively in cut rose flowers.

3.3 Object detection for GMD using YOLOv5

Among the methods employed for object detection, the YOLOv5 model demonstrated superior accuracy (mAP, precision, and recall) in comparison to the Faster R-CNN and SSD models (Supplementary Figure 6). Consequently, the YOLOv5 was chosen for object detection of GMD in cut roses in the present study. The object detection for GMD in cut roses was carried out by YOLOv5x models and the performance of the model was evaluated. The HSI

10.3389/fpls.2023.1296473

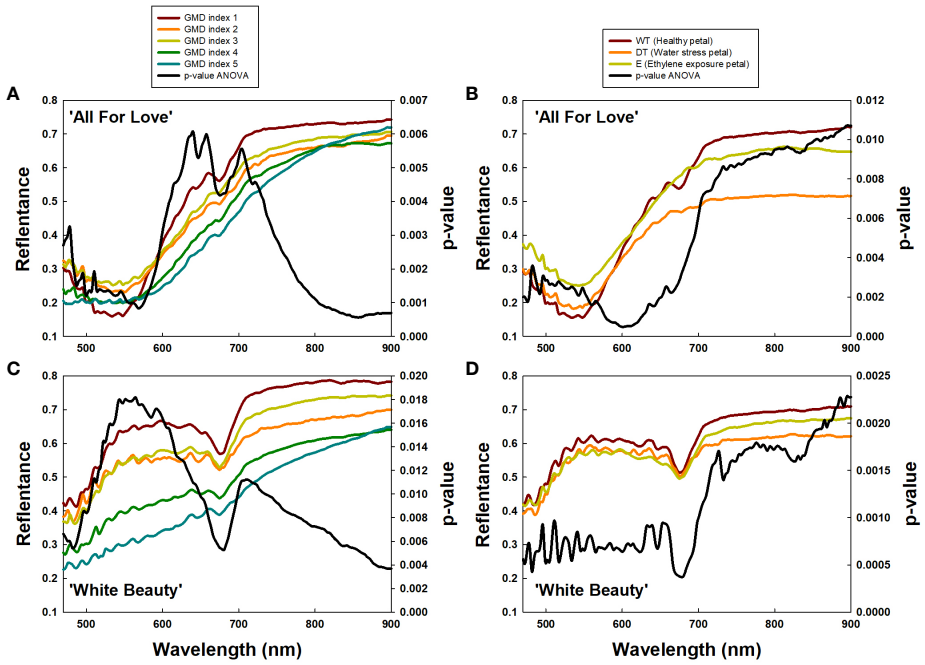


FIGURE 7
Average reflectance of wavelength (RW) of cut roses based on gray mold disease (GMD) index **(A, C)** and petal wilting level due to water stress or ethylene **(B, C)**, and *p*-values of a one-way ANOVA per RW **(A–D)**. GMD index was classified into five levels as 1, none; 2, slight symptoms (≤ 3%); 3, moderate symptoms (3-10%); 4, severe symptoms (11-50%); and 5, death of plants (> 50%). WT, wet transport; DT, dry transport; and E, ethylene exposure.

of cut roses was fed into the YOLOv5x model which was trained to identify the presence of GMD in petals. The model effectively detected small instances of GMD in rose petals (Figure 1C), demonstrating that YOLOv5x can predict the disease emergence at the early stages of the disease infection. The mAP represents the evaluation index of disease detection accuracy. In this study, mAP value was relatively high (82.1%) in 'All For Love' flowers (Figure 8A). The precision (86.2%) and recall (77.5%) values achieved by the model were also high in 'All For Love' flowers (Figure 8A). Whereas, the performance of the YOLOv5 model for 'White Beauty' flowers was slightly lower (mAP, 81.6%; precision, 85.1%; and recall, 78.4%) (Figure 8B). Nevertheless, these values were enough high and better than those of the prediction based on petal wilting levels (Supplementary Figure 7).

3.4 Prediction of vase life of cut roses using YOLOv5

The classification for vase life in cut roses was carried out by random forest models and the performance of the model was evaluated. The vase life of cut roses was trained into two

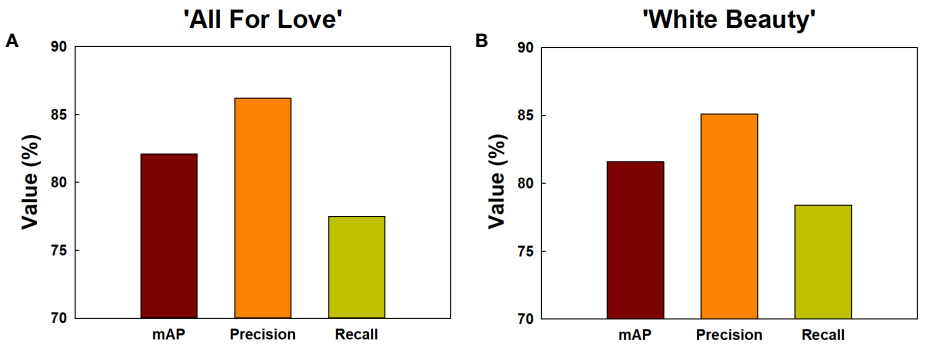


FIGURE 8
Detection and prediction of gray mold disease in cut roses 'All For Love' **(A)** and 'White Beauty' **(B)**. The performance of the prediction models by YOLOv5 was evaluated mAP, precision, and recall. mAP, the evaluation index of the detection accuracy; precision, the percentage of true positives (correctly detected objects) out of all the objects that is detected; recall, the percentage of true positives (correctly detected objects) out of all the objects that exist in the dataset.

categories as under 5 d (-5D) and over 5 d (+5D) based on the scores graded by the quality factors presented in Table 2. In this study, we evaluated the classification performance of the random forest algorithm in both cultivars. In 'All For Love' rose flowers, in the -5D case, the model displayed an F1 score of 89%, precision of 87%, and recall of 91% (Figure 9A). In contrast, in the +5D case, the performance was slightly lower (F1, 87%; precision, 85%; and Recall, 93%) (Figure 9C). In 'White Beauty' rose flowers, in the -5D case, the model yielded an F1 score of 85%, precision of 81%, and recall of 87% (Figure 9B). However, in the +5D case, the performance was slightly higher, with an F1 score of 88%, precision of 91%, and recall of 85% (Figure 9D).

The vase life prediction model was developed using YOLOv5x based on the detection of petal conditions (Figures 9E, F). As a result, the scatter plots showed a strong correlation between the predicted value and the observed value of the vase life evaluation (r^2 = 0.86 in 'All For Love' and 0.83 in 'White Beauty') (Figures 9E, F). This result indicates that the YOLOv5 model achieves a strong capacity for the vase life prediction of cut flowers by analyzing the large size of the complicated data obtained HSI.

4 Discussion

Postharvest conditions, such as dry transport, ethylene or high density of mold spores have been observed to decrease the longevity of the cut roses (Harkema et al., 2013; Ha et al., 2021; Ha et al., 2022). In this study, dry transport, ethylene exposure, and increased *B. cinerea* spore (induced by ethylene exposure and fungal conidial inoculation during transport simulation) significantly reduced vase life and positive water balance of cut roses. Dry transport, a practice involving storing cut flowers without water to facilitate transportation or control *B. cinerea* growth, can lead to dehydration and reduced vase life of cut flowers (Macnish et al.,

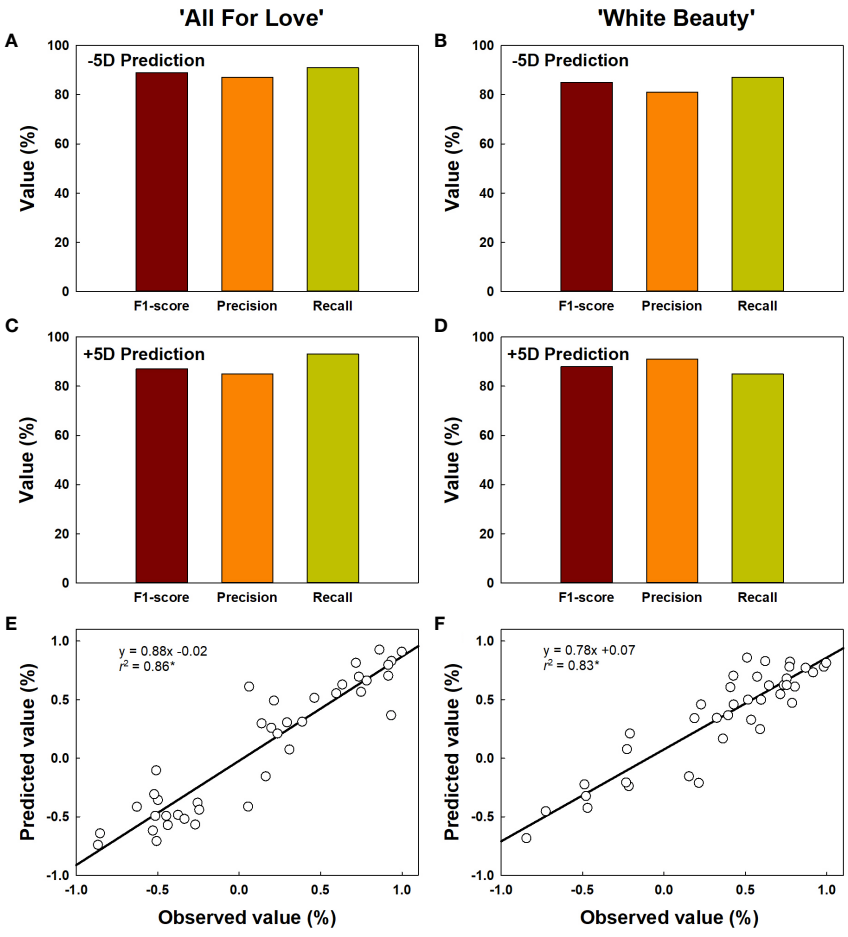


FIGURE 9
Prediction of vase life of cut roses 'All For Love' **(A, C)** and 'White Beauty' **(B, D)**. The performance of the prediction model by random forest was evaluated F1-score, precision, and recall. F1-score, the harmonic mean of precision; precision, the percentage of true positives (correctly detected objects) out of all the objects that is detected; recall, the percentage of true positives (correctly detected objects) out of all the objects that exist in the dataset. The accuracy of vase life prediction **(E, F)** by YOLOv5 and random forest. The vase life of cut roses were classed into two categories as over 5 d (+5D) and under 5 d (-5D) based on the scores graded by the quality factors in Table 2. The negative (-1.0-0) and positive (0.0-1.0) values by the linear regression analysis respectively indicate the probability that the vase life is -5D or +5D. Asterisk (*) represents a significant difference at p = 0.05 (n = 40).

2009; Harkema et al., 2013; Fanourakis et al., 2022). Ethylene exposure accelerates the senescence process of cut roses, leading to premature wilting, petal abscission, and overall reduced vase life (Macnish et al., 2010; In et al., 2017). During transportation, contact with *B. cinerea* spores or storage in conditions conductive to fungal growth can lead to infection, resulting in necrotic lesions and decay, ultimately reducing vase life of cut flowers (Ha et al., 2022).

Hyperspectral imaging is a non-contact method that analyses a wide range of light spectrums by scanning objects with hyperspectral cameras (Lowe et al., 2017; Cao et al., 2022; Xiang et al., 2022). The reflectance of light from plants at different wavelengths can be used to obtain information about various plant statuses and conditions (Sun J et al., 2021; Sukhova and Yudina, 2022). In this study, HSI of cut roses was used to observe distinct wavelength ranges of plants in various physiological states, such as GMD infection, water stress response, and senescence induction. The spectral reflectance at 470-680 nm was found to be strongly related to *B. cinerea* infection in the rose petals. The reflectance in this wavelength range is mainly affected by the absorption spectra of pigments in the leaves or petals of the plants (Rolfe and Scholes, 2010). *B. cinerea* infection would change the content and distribution of the pigment in the petals leading to changes in the spectral reflectance (López-López et al., 2016). While water stress causes changes in the water content of plant tissues which in turn affects the reflectance of light in the wavelength range of 700-900 nm (Elvanidi et al., 2018). Similarly, the reflectance at 700-900 nm was highly correlated to the petal wilting levels due to water stress after dry-transport or ethylene in cut flowers. Our results are consistent with those of previous studies showing that the reflectance at 400-680 nm is related to disease infection and the reflectance at higher 700 nm is sensitive to vegetation stress or water stress (Thenkabail et al., 2004; Köksal, 2011; Cao et al., 2022).

The YLOLv5x model was established to predict the potential incidence rate of GMD and the vase life of cut roses based on HSI data. We used the mAP@0.5 indicator to measure the prediction model's overall performance on the training test. The resulting mAP@0.5 value of the YOLOv5x model was approximately 80% in 'All For Love' roses, indicating that the model has a high prediction accuracy and can detect even small traces of fungal at early stages of disease development in rose petals. In previous studies, a similar detection performance was observed when YOLOv5 was used to predict powdery mildew disease and anthracnose in rubber plants (Chen et al., 2022). Our results also showed that the precision (78.6%) and recall (80.5%) values achieved from the model were also relatively high for 'All For Love' flowers, indicating that the model has a low chance of wrong detection (Qi et al., 2022). However, the disease detection performance was slightly lower for 'White Beauty' cultivar, possibly due to color similarity between white petals and GMD symptoms (Del Valle et al., 2018; Kasajima, 2019; Jiang et al., 2021).

Previously, the vase life prediction models in cut roses were developed by using the combination of machine learning and thermal imaging based on the different temperatures of petals among flower blooming stages (Choi and Lee, 2020). Evaluation of the flower quality of cut roses using a four-dimensional deep learning method was also studied based on the flower maturing status (Sun X et al., 2021). Despite the relatively good prediction accuracy of the models, an application of these techniques is difficult because the performance of the models is suitable only in limited conditions. In this study, the YOLOv5x models performed the vase life prediction well based on the detection of the flower states under different stress conditions and transportation methods. The results revealed that the models developed here are outstanding in the accuracy of the vase life prediction, consequently, applicability to the flower industry.

However, our model was developed using only two rose cultivars, thus further validation of the model with a larger dataset from various cultivars and environmental conditions is required to establish its general applicability. Furthermore, optimization of the YOLOv5 model, considering factors such as dataset size (Fang et al., 2021; Doherty et al., 2022) and computational resources (Junior and Ulson, 2021; Li et al., 2022), is crucial for improved performance and broader applicability.

5 Conclusion

In conclusion, our results have demonstrated the potential use of deep learning algorithms for detecting GMD and predicting the vase life of cut roses based on hyperspectral images of flower bud states. The finding from this study revealed that the spectral reflectance of 470 to 680 nm and 700 to 900 nm was closely related to GMD and plant physiological conditions, respectively in cut roses. The YOLOv5 model precisely detected and classified *B. cinerea* infection with high precision. The model also showed high predictive accuracy in evaluating the vase life of cut roses based on extensive image processing. With some modifications, the vase life prediction models developed in this study could be effective tools for constructing a flower longevity guarantee system for the flower industry.

Data availability statement

The datasets presented in this study can be found in online repositories. The names of the repository/repositories and accession number(s) can be found in the article/Supplementary Material.

Author contributions

Y-TK: Investigation, Data curation, Formal analysis, Software, Validation, Visualization, Writing – original draft. SH: Data curation, Formal analysis, Investigation, Writing – original draft. B-CI: Investigation, Conceptualization, Funding acquisition, Methodology, Project administration, Resources, Supervision, Writing – review & editing.

Funding

The author(s) declare financial support was received for the research, authorship, and/or publication of this article. This study was funded partially by Korea Institute of Planning and Evaluation for Technology in Food, Agriculture and Forestry (IPET) through Smart Agri Products Flow Storage Technology Development Program (322053-3), funded by Ministry of Agriculture, Food and Rural Affairs (MAFRA) and the Basic Research Program (NRF-2021R1I1A3A04037108) through the National Research Foundation of Korea (NRF) funded by Ministry of Education, Science and Technology.

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.

Supplementary material

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fpls.2023.1296473/full#supplementary-material

References

Ahmad, I., Yang, Y., Yue, Y., Ye, C., Hassan, M., Cheng, X., et al. (2022). Deep learning based detector YOLOv5 for identifying insect pests. *Appl. Sci.* 12, 10167. doi: 10.3390/app121910167

Amadi-Majd, M., Nejad, A. R., Mousavi-Fard, S., and Fanourakis, D. (2021). Deionized water as vase solution prolongs flower bud opening and vase life in carnation and rose through sustaining an improved water balance. *Eur. J. Hortic. Sci.* 86, 682–693. doi: 10.17660/eJHS.2021/86.6.12

Behmann, J., Steinrücken, J., and Plümer, L. (2014). Detection of early plant stress responses in hyperspectral images. *J. Photogramm. Remote Sens.* 93, 98–111. doi: 10.1016/j.isprsjprs.2014.03.016

Bulgari, R., Petrini, A., Cocetta, G., Nicoletto, C., Ertani, A., Sambo, P., et al. (2021). The impact of COVID-19 on horticulture: Critical issues and opportunities derived from an unexpected occurrence. *Horticulturae* 7, 124. doi: 10.3390/horticulturae7060124

Cao, Y., Yuan, P., Xu, H., Martínez-Ortega, J. F., 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. doi: 10.3389/fpls.2022.963170

Cape, J. N. (2003). Effects of airborne volatile organic compounds on plants. *Environ. Pollut.* 122, 145–157. doi: 10.1016/s0269-7491(02)00273-7

Chagué, V., Danit, L.V., Siewers, V., Schulze-Gronover, C., Tudzynski, P., Tudzynski, B., et al. (2006). Ethylene sensing and gene activation in Botrytis cinerea: a missing link in ethylene regulation of fungus-plant interactions? *Mol. Plant. Microbe. Interact.* 19, 33–42. doi: 10.1094/mpmi-19-0033

Chang, C., and Bleecker, A. B. (2004). Ethylene biology. More than a gas. *Plant Physiol.* 136, 2895–2899. doi: 10.1104/pp.104.900122

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

Choi, S. Y., and Lee, A. K. (2020). Development of a cut rose longevity prediction model using thermography and machine learning. *Hortic. Sci. Technol.* 38, 675–685. doi: 10.7235/HORT.20200061

Cravero, A., Pardo, S., Sepúlveda, S., and Muñoz, L. (2022). Challenges to use machine learning in agricultural big data: a systematic literature review. *Agronomy* 12, 748. doi: 10.3390/agronomy12030748

Del Valle, J. C., Gallardo-López, A., Buide, M. L., Whittall, J. B., and Narbona, E. (2018). Digital photography provides a fast, reliable, and noninvasive method to estimate anthocyanin pigment concentration in reproductive and vegetative plant tissues. *Eco. Evol.* 8, 3064–3076. doi: 10.1002/ece3.3804

Doherty, J., Gardiner, B., Kerr, E., Siddique, N., and Manvi, S. S. (2022). "Comparative study of activation functions and their impact on the YOLOv5 object detection model." in *International Conference on Pattern Recognition and Artificial Intelligence* (Cham: Springer International Publishing), 40–52.

Elvanidi, A., Katsoulas, N., Ferentinos, K. P., Bartzanas, T., and Kittas, C. (2018). Hyperspectral machine vision as a tool for water stress severity assessment in soilless tomato crop. *Biosyst. Eng.* 165, 25–35. doi: 10.1016/j.biosystemseng.2017.11.002

Fang, Y., Guo, X., Chen, K., Zhou, Z., and Ye, Q. (2021). Accurate and automated detection of surface knots on sawn timbers using YOLO-V5 model. *Bioresources* 16, 5390–5406. doi: 10.15376/biores.16.3.5390-5406

Fanourakis, D., Giday, H., Li, T., Kambourakis, E., Ligoxigakis, E. K., Papadimitriou, M., et al. (2016). Antitranspirant compounds alleviate the mild-desiccation-induced reduction of vase life in cut roses. *Postharvest Biol. Technol.* 117, 110–117. doi: 10.1016/j.postharvbio.2016.02.007

Fanourakis, D., Paparakis, V. M., Psyllakis, E., Tzanakakis, V. A., and Nektarios, P. (2022). The role of water relations and oxidative stress in the vase life response to prolonged storage: A case sttudy in chrysanthemum. *Agriculture.* 12 (2), 185. doi: 10.3390/agriculture12020185

Fanourakis, D., Pieruschka, R., Savvides, A., Macnish, A. J., Sarlikioti, V., and Woltering, E. J. (2013). Sources of vase life variation in cut roses: A review. *Postharvest Biol. Technol.* 78, 1–15. doi: 10.1016/j.postharvbio.2012.12.001

Fanourakis, D., Velez-Ramirez, A. I., In, B. C., Barendse, H., Van Meeteren, U., and Woltering, E. J. (2015). A Survey of preharvest conditions affecting the regulation of water loss during vase life. *Acta Hortic.* 1064, 195–204. doi: 10.17660/ActaHortic.2015.1064.22

Friedman, H., Agami, O., Vinokur, Y., Droby, S., Cohen, L., Refaeli, G., et al. (2010). Characterization of yield, sensitivity to *Botrytis cinerea* and antioxidant content of several rose species suitable for edible flowers. *Sci. Hortic.* 123, 395–401. doi: 10.1016/j.scienta.2009.09.019

Gabellini, S., and Scaramuzzi, S. (2022). Evolving consumption trends, marketing strategies, and governance settings in ornamental horticulture: A grey literature review. *Horticulturae* 8, 234. doi: 10.3390/horticulturae8030234

Guo, Y., Liu, Y., Oerlemans, A., Lao, S., Wu, S., and Lew, M. S. (2016). Deep learning for visual understanding: A review. *Neurocomputing* 187, 27–48. doi: 10.1016/j.neucom.2015.09.116

Ha, S. T. T., Kim, Y.-T., Yeam, I., Choi, H. W., and In, B.-C. (2022). Molecular dissection of rose and *Botrytis cinerea* pathosystems affected by ethylene. *Postharvest Biol. Technol.* 194, 112104. doi: 10.1016/j.postharvbio.2022.112104

Ha, S. T. T., Nguyen, T. K., and Lim, J. H. (2021). Effects of air-exposure time on water relations, longevity, and aquaporin-related gene expression of cut roses. *Hortic. Environ. Biotechnol.* 62, 63–75. doi: 10.1007/s13580-020-00302-1

Harkema, H., Mensink, M. G. J., Somhorst, D. P. M., Pedreschi, R. P., and Westra, E. H. (2013). Reduction of *Botrytis cinerea* incidence in cut roses (*Rosa hybrida* L.) during long term transport in dry conditions. *Postharvest Biol. Technol.* 76, 135–138. doi: 10.1016/j.postharvbio.2012.10.003

In, B.-C., Ha, S. T. T., Lee, Y. S., and Lim, J. H. (2017). Relationships between the longevity, water relations, ethylene sensitivity, and gene expression of cut roses. *Postharvest Biol. Technol.* 131, 74–83. doi: 10.1016/j.postharvbio.2017.05.003

In, B.-C., Inamoto, K., and Doi, M. (2009). A neural network technique to develop a vase life prediction model of cut roses. *Postharvest Biol.Technol.* 52, 273–278. doi: 10.1016/j.postharvbio.2009.01.001

In, B.-C., Inamoto, K., Doi, M., and Park, S.-A. (2016a). Using thermography to estimate leaf transpiration rates in cut roses for the development of vase life prediction models. *Hortic. Environ. Biotechnol.* 57, 53–60. doi: 10.1007/s13580-016-0117-6

In, B.-C., Lee, J.-H., Lee, A.-K., and Lim, J. H. (2016b). Conditions during export affect the potential vase life of cut roses (*Rosa hybrida* L.). *Hortic. Environ. Biotechnol.* 57, 504–510. doi: 10.1007/s13580-016-1119-0

In, B.-C., and Lim, J. H. (2018). Potential vase life of cut roses: Seasonal variation and relationships with growth conditions, phenotypes, and gene expressions. *Postharvest Biol. Technol.* 135, 93–103. doi: 10.1016/j.postharvbio.2017.09.006

Ji, Y., Sun, L., Li, Y., and Ye, D. (2019). Detection of bruised potatoes using hyperspectral imaging technique based on discrete wavelet transform. *Infrared Phys. Technol.* 103, 103054. doi: 10.1016/j.infrared.2019.10305

Jiang, Q., Wu, G., Tian, C., Li, N., Yang, H., Bai, Y., et al. (2021). Hyperspectral imaging for early identification of strawberry leaves diseases with machine learning and spectral fingerprint features. *Infrared Phys. Technol.* 118, 103898. doi: 10.1016/j.infrared.2021.10389

Junior, L. C. M., and Ulson, J. (2021). "Real time weed detection using computer vision and deep learning," in *2021 14th IEEE International Conference on Industry Applications (INDUSCON)*, São Paulo, Brazil, pp. 1131–1137. doi: 10.1109/INDUSCON51756.2021.9529761

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

Kasajima, I. (2019). Measuring plant colors. *Plant Biotechnol.* 36, 63–75. doi: 10.5511/plantbiotechnology.19.0322a

Köksal, E. S. (2011). Hyperspectral reflectance data processing through cluster and principal component analysis for estimating irrigation and yield related indicators. *Agric. Water Manage.* 98, 1317–1328. doi: 10.1016/j.agwat.2011.03.014

Lay, L., Lee, H. S., Tayade, R., Ghimire, A., Chung, Y. S., Yoon, Y., et al. (2023). Evaluation of soybean wildfire prediction via hyperspectral imaging. *Plants* 12, 901. doi: 10.3390/plants12040901

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

Li, J., Qiao, Y., Liu, S., Zhang, J., Yang, Z., and Wang, M. (2022). An improved YOLOv5-based vegetable disease detection method. *Comput. Electron. Agr.* 202, 107345. doi: 10.1016/j.compag.2022.107345

Liu, D., Zeng, X. A., and Sun, D. W. (2015). Recent developments and applications of hyperspectral imaging for quality evaluation of agricultural products: a review. *Crit. Rev. Food. Sci. Nutr.* 55, 1744–1757. doi: 10.1080/10408398.2013.777020

López-López, M., Calderón, R., González-Dugo, V., Zarco-Tejada, P. J., and Fereres, E. (2016). Early detection and quantification of almond red leaf blotch using high-resolution hyperspectral and thermal imagery. *Remote Sens.* 8, 276. doi: 10.3390/rs8040276

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, 80. doi: 10.1186/s13007-017-0233-z

Macnish, A. J., De Theije, A., Reid, M. S., and Jiang, C. Z. (2009). An alternative postharvest handing strategy for cut flowers-dry handling after harvest. *Acta Hortic.* 847, 215–222. doi: 10.17660/ActaHortic.2009.847.27

Macnish, A. J., Leonard, R. T., Borda, A. M., and Nell, T. A. (2010). Genotypic variation in the postharvest performance and ethylene sensitivity of cut rose flowers. *J. Am. Soc Hortic. Sci.* 45, 790–796. doi: 10.21273/HORTSCI.45.5.790

Martínez-Romero, D., Guillén, F., Castillo, S., Zapata, P. J., Serrano, M., and Valero, D. (2009). Development of a carbon-heat hybrid ethylene scrubber for fresh horticultural produce storage purposes. *Postharvest Biol. Technol.* 51, 200–205. doi: 10.1016/j.postharvestbio.2008.07.013

Mo, C., Kim, G., Lim, J., Kim, M. S., Cho, H., and Cho, B.-K. (2015). Detection of lettuce discoloration using hyperspectral reflectance imaging. *Sensors* 15, 29511–29534. doi: 10.3390/s151129511

Nasiri, A., Taheri-Garavand, A., Fanourakis, D., Zhang, Y.-D., and Nikoloudakis, N. (2021). Automated grapevine cultivar identification via leaf imaging and deep convolutional neural networks: A proof-of-concept study employing primary Iranian varieties. *Plants* 10, 1628. doi: 10.3390/plants10081628

Qi, J., Liu, X., Liu, K., Xu, F., Guo, H., Tian, X., et al. (2022). An improved YOLOv5 model based on visual attention mechanism: Application to recognition of tomato virus disease. *Comput. Electron. Agr.* 194, 106780. doi: 10.1016/j.compag.2022.106780

Rahman, A., Kandpal, L. M., Lohumi, S., Kim, M. S., Lee, H., Mo, C., et al. (2017). Nondestructive estimation of moisture content, pH and soluble solid contents in intact tomatoes using hyperspectral imaging. *Appl. Sci.* 7, 109. doi: 10.3390/app7010109

Ramamoorthy, P., Samiappan, S., Wubben, M. J., Brooks, J. P., Shrestha, A., Panda, R. M., et al. (2022). Hyperspectral reflectance and machine learning approaches for the detection of drought and root-knot nematode infestation in cotton. *Remote Sens.* 14, 4021. doi: 10.3390/rs14164021

Redmon, J., Divvala, S., Girshick, R., and Farhadi, A. (2016). "You Only Look Once: unified, real-time object detection," in *2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, Las Vegas, NV, USA, pp. 779–788. doi: 10.1109/CVPR.2016.91

Reid, M. S., Mokhtari, M., Lieth, J. H., van Doorn, W. G., and Evans, R. Y. (1996). Modeling the postharvest life of cut roses. *Acta Hortic.* 424, 137–144. doi: 10.17660/ActaHortic.1996.424.24

Rolfe, S. A., and Scholes, J. D. (2010). Chlorophyll fluorescence imaging of plant–pathogen interactions. *Protoplasma* 247, 163–175. doi: 10.1007/s00709-010-0203-z

Staby, G. L., and Cunningham, M. S. (1980). Predicting longevity of carnations to reduce postharvest shrinkage. *Ohio Rep.* 65, 54–55.

Stead, A. D., Gay, A., Taylor, J., Ougham, H., Wagstaff, C., and Rogers, H. J. (2018). Hyperspectral imaging as a means to assess quality issues of cut flowers. *Acta Hortic.* 1263, 359–366. doi: 10.17660/ActaHortic.2009.1263.47

Suarez, M. B., Walsh, K., Boonham, N., O'Neill, T., Pearson, S., Barker, I., et al. (2005). Development of real-time PCR (TaqMan) assays for the detection and quantification of *Botrytis cinerea* in planta. *Plant Physiol. Biochem.* 43, 890–899. doi: 10.1016/j.plaphy.2005.07.003

Sukhova, E., and Yudina, L. (2022). Modified photochemical reflectance indices as new tool for revealing Influence of drought and heat on pea and wheat plants. *Plants* 11, 1380. doi: 10.3390/plants11101308

Sun, X., Li, Z., Zhu, T., and Ni, C. (2021). Four-dimension deep learning method for flower quality grading with depth information. *Electronics* 10, 2353. doi: 10.3390/electronics10192353

Sun, J., Yang, L., Yang, X., Wei, J., Li, L., Guo, E., et al. (2021). Using spectral reflectance to estimate the leaf chlorophyll content of maize inoculated with *Arbuscular Mycorrhizal* fungi under water stress. *Front. Plant Sci.* 12. doi: 10.3389/fpls.2021.646173

Susič, N., Žibrat, U., Širca, S., Strajnar, P., Razinger, J., Knapič, M., et al. (2018). Discrimination between abiotic and biotic drought stress in tomatoes using hyperspectral imaging. *Sens. Actuators B Chem.* 273, 842–852. doi: 10.1016/j.snb.2018.06.121

Taghizadeh, M., Gowen, A. A., and O'donnell, C. P. (2011). The potential of visible-near infrared hyperspectral imaging to discriminate between casing soil, enzymatic browning and undamaged tissue on mushroom (*Agaricus bisporus*) surfaces. *Comput. Electron. Agr.* 77, 74–80. doi: 10.1016/j.compag.2011.03.010

Taheri-Garavand, A., Mumivand, H., Fanourakis, D., Fatahi, S., and Taghipour, S. (2021). An artificial neural network approach for non-invasive estimation of essential oil content and composition through considering drying processing factors: A case study in Mentha aquatica. *Ind. Crop Prod.* 171, 113985. doi: 10.1016/j.indcrop.2021.113985

Thenkabail, P. S., Enclona, E. A., Ashton, M. S., and van der Meer, B. (2004). Accuracy assessments of hyperspectral waveband performance for vegetation analysis applications. *Remote Sens. Environ.* 91, 354–376. doi: 10.1016/j.rse.2004.03.013

Tian, H., Wang, T., Liu, Y., Qiao, X., and Li, Y. (2020). Computer vision technology in agricultural automation —A review. *Inf. Process. Agric.* 7, 1–19. doi: 10.1016/j.inpa.2019.09.006

Tromp, S.-O., van der Sman, R. G. M., Vollebregt, H. M., and Woltering, E. J. (2012). On the prediction of the remaining vase life of cut roses. *Postharvest Biol. Technol.* 70, 42–50. doi: 10.1016/j.postharvbio.2012.04.003

VBN (2014). *Evaluation cards for Rosa* (The Netherlands: FloraHollandAalsmeer).

Vehniwal, S. S., and Abbey, L. (2019). Cut flower vase life – influential factors, metabolism, and organic formulation. *Horticult. Int. J.* 3, 275–281. doi: 10.15406/hij.2019.03.00142

Veys, C., Chatziavgerinos, F., Alsuwaidi, A., Hibbert, J., Hansen, M., Bernotas, G., et al. (2019). Multispectral imaging for presymptomatic analysis of light leaf spot in oilseed rape. *Plant Methods* 15, 4. doi: 10.1186/s13007-019-0389-9

Wang, K. L.-C., Li, H., and Ecker, J. R. (2002). Ethylene biosynthesis and signaling networks. *Plant Cell* 14, S131–S151. doi: 10.1105/tpc.001768

Wieme, J., Mollazade, K., Malounas, I., Zude-Sasse, M., Zhao, M., Gowen, A., et al. (2022). Application of hyperspectral imaging systems and artificial intelligence for quality assessment of fruit, vegetables and mushrooms: A review. *Biosyst. Eng.* 222, 156–176. doi: 10.1016/j.biosystemseng.2022.07.013

Williamson, B., Tudzynski, B., Tudzynski, P., and Van Kan, J. A. (2007). *Botrytis cinerea*: the cause of grey mould disease. *Mol. Plant Pathol.* 8, 561–580. doi: 10.1111/j.1364-3703.2007.00417.x

Xiang, Y., Chen, Q., Su, Z., Zhang, L., Chen, Z., Zhou, G., et al. (2022). Deep learning and hyperspectral images based tomato soluble solids content and firmness estimation. *Front. Plant Sci.* 13. doi: 10.3389/fpls.2022.860656

Xue, J., Yang, F., and Gao, J. (2009). Isolation of *Rh-TIP1;1*, an aquaporin gene and its expression in rose flowers in response to ethylene and water deficit. *Postharvest Biol. Technol.* 51, 407–413. doi: 10.1016/j.postharvbio.2008.08.011

Yao, J., Qi, J., Zhang, J., Shao, H., Yang, J., and Li, X. (2021). A real-time detection algorithm for kiwi fruit defects based on YOLOv5. *Electronics* 10, 1711. doi: 10.3390/electronics10141711

Zhang, C., Guo, C., Liu, F., Kong, W., He, Y., and Lou, B. (2016). Hyperspectral imaging analysis for ripeness evaluation of strawberry with support vector machine. *J. Food Eng.* 179, 11–18. doi: 10.1016/j.jfoodeng.2016.01.002

Zhang, Y., He, S., Wa, S., Zong, Z., and Liu, Y. (2021). Using generative module and pruning inference for the fast and accurate detection of apple flower in natural environments. *Information* 12, 495. doi: 10.3390/info12120495

frontiers | Frontiers in Plant Science

TYPE Original Research
PUBLISHED 15 January 2024
DOI 10.3389/fpls.2023.1342970

OPEN ACCESS

EDITED BY
Lie Deng,
Southwest University, China

REVIEWED BY
Zheli Wang,
China Agricultural University, China
Yong Hao,
East China Jiaotong University, China

*CORRESPONDENCE
Yuanyuan Song
yyuansong@fafu.edu.cn
Li Gu
guli5101@163.com

†These authors have contributed equally to this work

RECEIVED 22 November 2023
ACCEPTED 27 December 2023
PUBLISHED 15 January 2024

CITATION
Zhang T, Lu L, Song Y, Yang M, Li J, Yuan J, Lin Y, Shi X, Li M, Yuan X, Zhang Z, Zeng R, Song Y and Gu L (2024) Non-destructive identification of *Pseudostellaria heterophylla* from different geographical origins by Vis/NIR and SWIR hyperspectral imaging techniques. *Front. Plant Sci.* 14:1342970. doi: 10.3389/fpls.2023.1342970

Non-destructive identification of *Pseudostellaria heterophylla* from different geographical origins by Vis/NIR and SWIR hyperspectral imaging techniques

Tingting Zhang[1,2†], Long Lu[1,2†], Yihu Song[1†], Minyu Yang[1], Jing Li[1], Jiduan Yuan[3], Yuquan Lin[4], Xingren Shi[4], Mingjie Li[1,2], Xiaotan Yuan[3], Zhongyi Zhang[1,2], Rensen Zeng[1,2], Yuanyuan Song[1,2*] and Li Gu[1,2*]

[1]Key Laboratory of Ministry of Education for Genetics, Breeding and Multiple Utilization of Crops, College of Agriculture, Fujian Agriculture and Forestry University, Fuzhou, China, [2]Key Laboratory of Biological Breeding for Fujian and Taiwan Crops, Ministry of Agriculture and Rural Affairs, Fujian Agriculture and Forestry University, Fuzhou, China, [3]Pharmaceutical Development Board of Zherong County, Ningde, China, [4]Huzhou Wuxing Jinnong Ecological Agriculture Development Co., Ltd, Huzhou, China

The composition of *Pseudostellaria heterophylla* (Tai-Zi-Shen, TZS) is greatly influenced by the growing area of the plants, making it significant to distinguish the origins of TZS. However, traditional methods for TZS origin identification are time-consuming, laborious, and destructive. To address this, two or three TZS accessions were selected from four different regions of China, with each of these resources including distinct quality grades of TZS samples. The visible near-infrared (Vis/NIR) and short-wave infrared (SWIR) hyperspectral information from these samples were then collected. Fast and high-precision methods to identify the origins of TZS were developed by combining various preprocessing algorithms, feature band extraction algorithms (CARS and SPA), traditional two-stage machine learning classifiers (PLS-DA, SVM, and RF), and an end-to-end deep learning classifier (DCNN). Specifically, SWIR hyperspectral information outperformed Vis/NIR hyperspectral information in detecting geographic origins of TZS. The SPA algorithm proved particularly effective in extracting SWIR information that was highly correlated with the origins of TZS. The corresponding FD-SPA-SVM model reduced the number of bands by 77.2% and improved the model accuracy from 97.6% to 98.1% compared to the full-band FD-SVM model. Overall, two sets of fast and high-precision models, SWIR-FD-SPA-SVM and SWIR-FD-DCNN, were established, achieving accuracies of 98.1% and 98.7% respectively. This work provides a potentially efficient alternative for rapidly detecting the origins of TZS during actual production.

KEYWORDS

***Pseudostellaria heterophylla*, geographical origin, hyperspectral imaging, machine learning, deep learning**

1 Introduction

Pseudostellaria heterophylla (Miq.) Pax, also known as Tai-zi-shen (TZS), is a perennial herbaceous plant belonging to the *Caryophyllaceae* family (Li et al., 2016). Its roots have a long history of use as medicinal and edible plants in Asian countries, including China and Korea. This plant is renowned for its safety and its content of polysaccharides, saponins, cyclopeptides, sterols, oils, and other volatile oily substances, which contribute to enhancing the human immune system (Wong et al., 1994). TZS is commonly employed as a substitute for ginseng and American ginseng, addressing issues such as loss of appetite and serving as a potent tonic. Wild TZS resources are widely distributed in various provinces of China, such as Fujian, Guizhou, Jiangsu, and Anhui. However, the composition of TZS varies among different geographical origins. To evaluate TZS quality based on polysaccharide and saponin content, it is crucial to consider cultivated plants from specific regions, such as Jiangsu Province and Fujian Province (Shi et al., 2013). Therefore, distinguishing the origins of TZS is significant.

Most traditional methods used to identify the origins and grades of herbs rely on external characteristics such as shape, color, microstructure, and odor. However, the similarity of the external features of TZS makes it difficult to detect their origins and grades, especially after processing (Wu et al., 2018). Currently, the identification of TZS is conducted through techniques like High-Performance Liquid Chromatography, Gas Chromatography-Mass Spectrometry, ninhydrin color, and other analytical methods that aim to detect specific active components (Lin et al., 2011). However, these methods are time-consuming, labor-intensive, expensive, and require the use of large quantities of organic solvents, which can potentially harm the environment. Thus, there is an urgent need for a rapid and accurate analytical method to determine the origins of TZS.

Today, hyperspectral imaging (HSI) is widely utilized in agri-food product quality and safety control (Lu et al., 2020; Tian et al., 2023). The HSI combines conventional imaging and spectroscopic techniques to present a hypercube with one spectral dimension and two spatial dimensions. This allows it to provide both spatial and spectral information about the sample (Zareef et al., 2021). This information is closely related to the chemical composition and physical properties of the sample (Delwiche and Kim, 2000). Therefore, the HSI technique has attracted considerable attention in distinguishing between similar groups of biological materials such as maize (Wang et al., 2022), wheat (Zhang et al., 2018; Zhang et al., 2022b), wolfberries (Zhang et al., 2020a; Dong et al., 2022).

Artificial intelligence technology has assumed a crucial role in numerous global domains. Machine learning (ML) is an essential approach in studying artificial intelligence. In recent years, the ML field has experienced a significant transformation owing to the development of novel deep learning (DL) classifiers. DL, with its capacity to comprehend intricate and representative patterns from vast datasets, has found applications across diverse domains. Shallow Convolutional Neural Networks (CNN), representative algorithms for DL, have been proven by previous studies to be ideal for analyzing and processing HSI data (Liu et al., 2020; Zhang et al., 2022a). The complexity of traditional neural networks is reduced by a simple network structure. The "end-to-end" design concept, coupled with the hidden-neuron network structure, empowers us to autonomously extract relevant data features and optimize large datasets for accurate target classification (Fu et al., 2018). To the best of the researcher's knowledge, the combination of HSI and DL algorithms to recognize the geographical origins and quality grades of TZS has not been reported yet.

Therefore, in this study, we utilized HIS combined with DL and ML techniques for the evaluation of the geographical origins of *Pseudostellaria heterophylla* (Miq.) Pax (TZS). The successive projection algorithm (SPA) and competitive adaptive weighted sampling (CARS) were employed to extract spectral information that is highly correlated with the origins of TZS. ML methods, such as partial least squares discriminant analysis (PLS-DA), random forests (RF), and support vector machines (SVMs), were also compared as modeling approaches alongside deep convolutional neural network (DCNN) architectures.

2 Materials and methods

2.1 Sample preparation

The TZS samples utilized in this study were collected from four distinct geographical regions in China, including Guizhou, Fujian, Jiangsu, and Anhui Provinces. To improve the applicability of the model, we attempted to increase the complexity of the sample composition. Two or three germplasm resources for each geographical region were selected for this reason, encompassing different quality grades of TZS (Table 1). We randomly selected 3249 samples from the TZS accessions, covering all three quality grades. The quality grades of TZS samples were determined according to the commercial grades for Chinese materia medica-PSEUDOSTELLARIAE RADIX (T/CACM 1021.127-2018). Specifically, the first-grade samples were characterized by roughly straight shapes, with diameters of the thickened root section greater than or equal to 0.4 cm and individual weights greater than or equal to 0.4 g. Similarly, the second-grade samples also had roughly straight shapes, with diameters of the thickened root section greater than or equal to 0.3 cm and individual weights greater than or equal to 0.2 g. In contrast, the third-grade samples were classified as bent, with diameters of the thickened root section below 0.3 cm and individual weights below 0.2 g. Additionally, to capture comprehensive spectral information of the TZS, both sides of each sample were scanned using visible and near-infrared (Vis/NIR) as well as shortwave infrared (SWIR) hyperspectral instruments. Consequently, a total of 12996 hyperspectral images of the TZS samples were obtained. Typical TZS images from different origins and quality grades are shown in Figure 1.

2.2 Hyperspectral image acquisition and correction

The hyperspectral imaging (HSI) system comprised instruments for both Vis/NIR and SWIR spectral ranges. The Vis/

10.3389/fpls.2023.1342970

TABLE 1 The geographical origins and quality grades of TZS.

Geographical origins	Name	Year	Number of samples in different quality grades		
			1	2	3
Fujian province	Zheshen 1	2022	96	82	80
	Zheshen 4	2022	82	84	84
	Landrace 1	2023	104	104	102
Guizhou province	Guishen 1	2023	92	92	92
	Landrace 1	2023	96	88	90
	Landrace 2	2023	81	90	92
Jiangsu province	Landrace 1	2023	132	132	132
	Landrace 2	2023	132	136	134
Anhui province	Xuanshen 1	2023	92	90	92
	Xuanshen 2	2023	92	90	90
	Xuanshen 3	2023	92	90	92

NIR instrument consists of a GaiaField Pro-V10E spectrometer (Specim, Spectral Imaging Ltd., Finland), a high-resolution camera (Sichuan Dualix Spectral Imaging Technology Co., Ltd., China), and two 150 W halogen light sources. The SWIR instrument is composed of a GaiaField Pro-N17E-HR spectrometer (Specim, Spectral Imaging Ltd., Finland), a shortwave infrared high-resolution camera (Sichuan Dualix Spectral Imaging Technology Co., Ltd., China), and two 150 W tungsten halogen lamps. These two instruments employed a sample stage that was electrically positioned and controlled by a stepper motor. A computer equipped with SpecView Software (Sichuan Dualix Spectral Imaging Technology Co., Ltd., China) was provided. The

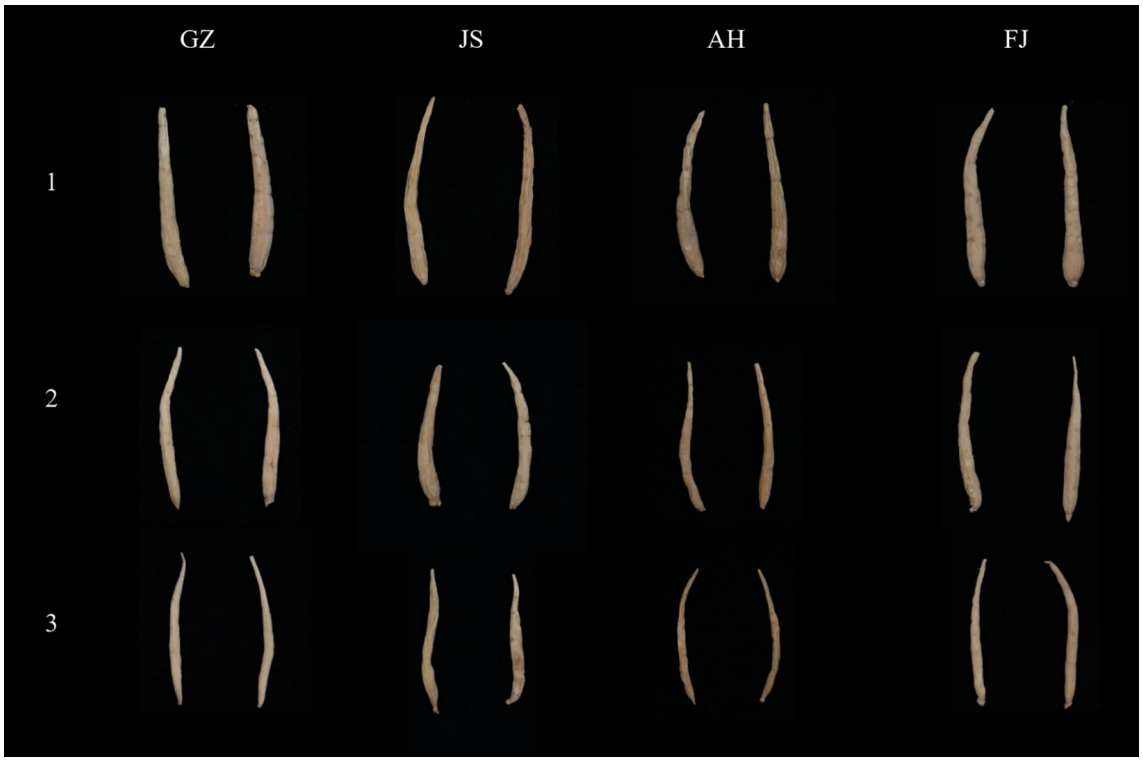


FIGURE 1
Typical TZS samples from Guizhou (GZ), Jiangsu (JS), Anhui (AH) and Fujian (FJ) Provinces. The numbers "1, 2, 3" on the left represent the different quality levels of TZS.

instruments mentioned above were enclosed within a box featuring a black inner surface, thus constituting the HSI system.

To eliminate errors like baseline drift, the HSI system should be preheated by powering it on for 30 minutes before image collection. Subsequently, non-deformed three-dimensional hyperspectral images (x, y, λ), commonly known as "hypercubes", were obtained by placing the TZS samples on the platform. To obtain high-quality hyperspectral reflectance images of the TZS samples, the Vis/NIR-HSI image acquisition parameters of motor speed, exposure time, and object distance were set at 1.18 cm/s, 7.2 ms, and 25 cm through several attempts. Similarly, the SWIR-HSI image acquisition parameters of motor speed, exposure time, and object distance were adjusted to 1.5 cm/s, 4.5 ms, and 25 cm after several attempts. Hyperspectral image correction was then conducted by using white and black references according to the method depicted in Zhang et al. (2020b).

2.3 Spectral data extraction

Each TZS sample was then considered as a region of interest (ROI) and was segmented from the black background by threshold segmentation. The difference between the sample and the background reflectance was maximum at the 801.05 nm band (Vis/NIR) and 1005.67 nm band (SWIR), so every sample was segmented at these bands separately. The spectra of pixels belonging to the same TZS sample were averaged to derive average spectra, which were then utilized for discrimination analysis purposes. The head and tail bands were eliminated from the spectra to minimize the effects of instability stemming from random noise. Consequently, 673 bands from 400.20 nm to 999.75 nm for the Vis/NIR and 482 bands from 900.96 nm to 1700.43 nm for the SWIR hypercubes were utilized for future analysis.

2.4 Spectral data preprocessing

To minimize the potential effects of overlapping or light noise across different spectral wavelengths (Alchanatis et al., 2005), as well as to assess the impact of various pre-processing methods on the classification of TZS origins, several spectral pre-processing techniques were investigated and applied to the raw spectra. The evaluated pre-processing techniques included standard normal variate (SNV) (Barnes et al., 1989), Detrend (DT), and Savitzky–Golay first derivative (FD) (Zhang et al., 2020b).

2.5 Multivariate data analysis

In this study, various machine learning algorithms were employed, including traditional two-stage methods such as partial least squares discriminant analysis (PLS-DA), support vector machine (SVM), and random forest (RF), as well as an end-to-end deep learning algorithm known as the deep convolutional neural network (DCNN). The purpose of these algorithms was to distinguish TZS samples into different origin groups.

2.5.1 Traditional two-stage machine learning algorithms

PLS-DA is a widely practiced classifier that is considered a supervised technique that maximizes the distinction between different groups of samples (Nie et al., 2019). RF is an ensemble learning algorithm developed by Leo Breiman, inspired by the earlier work of Amit and Geman (Breiman, 2001). RF offers several advantages, including fast training speed, few tuning parameters, and the ability to handle high-dimensional data effectively. At its core, RF is a collection of decision trees that work together to make predictions (Tian et al., 2021). SVM is a non-probabilistic, linear, binary classifier used to classify linear and nonlinear data by learning a hyperplane. In nonlinear classification, SVM uses a kernel function to map original data to high-dimensional data and build hyperplanes to optimally classify the closest training samples in different classes (Burges and discovery, 1998; Wang et al., 2023). In this study, the radial basis function (RBF) kernel was selected for the SVM algorithm and the penalty coefficient c and kernel parameter g were optimized by a grid search procedure in the range of 2^{-8}–2^{8} through five-fold cross-validation.

2.5.2 Deep learning algorithms

The DCNN is a widely recognized deep learning architecture that is inspired by the visual perception mechanisms found in living organisms (Zhang et al., 2019). A one-dimensional DCNN was developed as the primary classifier to process the data from each source. The DCNN architecture consisted of three convolutional blocks, one flattening layer, and five fully-connected layers. Each convolutional block comprised convolutional, batch normalization, maximum pooling, and dropout layers. To extract local features from the spectral information effectively, while reducing its dimensions and enhancing non-linearity, we utilized 1×3 convolution kernels with a stride and padding of 1 (Yu et al., 2021). The first and second convolutional layers were configured with 32, 64, and 128 filters, respectively.

Utilizing rectified linear units (ReLUs) in the DCNN resulted in faster training and helped mitigate model overfitting compared to networks using older units (Nie et al., 2019). To facilitate learning and reduce the emphasis on initialization, batch normalization was applied before each dense layer and after each convolutional layer (Ioffe and Szegedy, 2015). The fully connected (Fc) layers were composed of 512, 128, 64, 32, and 4 neurons, respectively. To convert the DCNN output into probabilities for each category, a softmax function was introduced to the activation function of Fc5 (Kumar et al., 2022). The categorical cross-entropy loss function was employed to measure the distance between the probability distribution of the DCNN output values and the true values. To minimize the loss function, we utilized an adaptive moment estimation algorithm with a learning rate of 0.001, beta_1 of 0.9, and beta_2 of 0.99 (Yu et al., 2021).

2.6 Model implementation and evaluation

The sample data were randomly divided into training and validation sets in a ratio of 7: 3. The classification accuracy, which

was used to evaluate the performance of the models, was determined by calculating the ratio of correctly classified samples to the total number of samples. The diagrams were developed using Origin Pro 9.0 (Origin Lab Corporation, Northampton, MA, USA). MATLAB R2019b (The MathWorks, Natick, MA, USA) was utilized for spectrum extraction, spectrum preprocessing, and ML model development. The DCNN model was built using Keras, a renowned deep-learning framework (https://keras.io/zh/).

3 Results and discussions

3.1 Reflectance spectral characteristics of the samples

The raw reflectance spectra of all the TZS samples from different origins were presented in Figure 2, covering the spectral ranges of 400-1000 nm and 1000-2000 nm. For the same spectral range, the TZS from different origins exhibited similar trends in general. This was similar to the previous researches conducted on discriminating maize varieties, determining the geographical origin of wolfberries, and assessing the quality of potatoes (López-Maestresalas et al., 2016; Dong et al., 2022). Although the spectral curves exhibited similar trends across various origins, there were variations in reflectance intensities. This discrepancy suggested that while the types of internal substances were similar, their concentrations differed among the different origins.

3.2 Preliminary principal component analysis to explore natural clustering of TZS samples

Two Principal Component Analysis (PCA) models were initially developed using the Vis/NIR and SWIR spectra of the TZS samples with the aim of observing the initial structure of the data from different geographically originated samples and detecting any anomalies among the samples. Three principal components (PCs) were selected for the Vis/NIR range, which accounted for 98.4% of the total variance (Figure 3A). Similarly, three PCs were chosen for the SWIR range, explaining 99.0% of the total variance (Figure 3B). According to the analysis, significant overlap between samples from different origins was observed in both spectral regions. It was worth noting that the distribution patterns of samples from different origins between the two spectral regions varied to some extent. In the Vis/NIR range, the samples of TZS from Jiangsu origin were slightly separated from the samples of other origins. Yet, this was not observed in the SWIR region. These differences provided a theoretical basis for further in-depth mining of the data in the two spectral regions separately.

3.3 Classification models based on full wavelengths

The PLS-DA, RF, SVM, and DCNN classification models were constructed by combining the SNV, DT and FD algorithms based

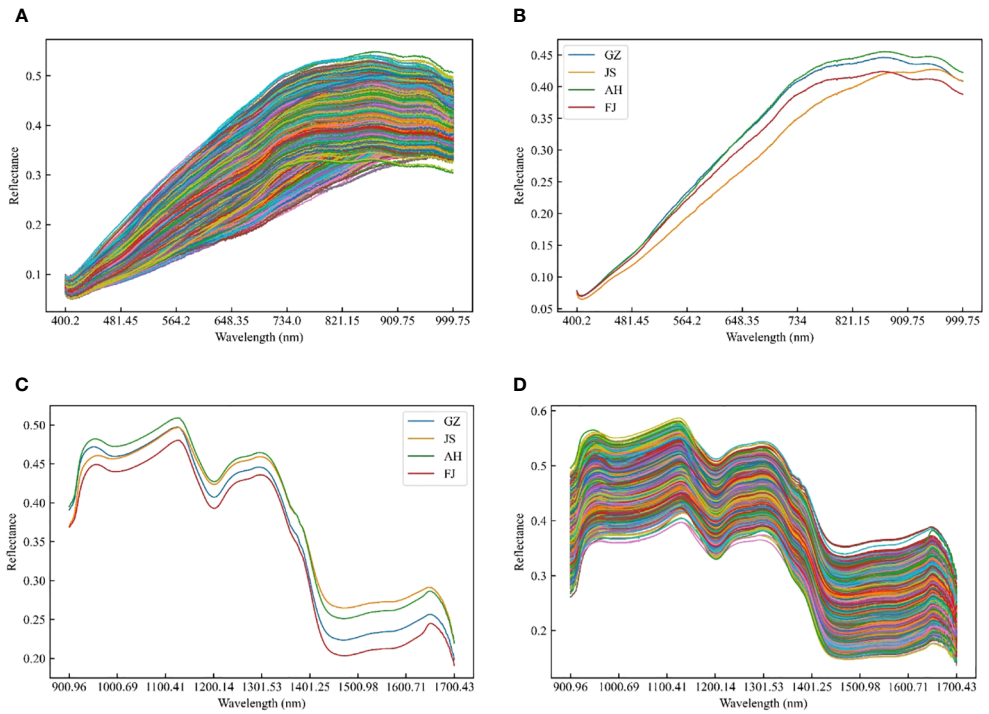


FIGURE 2
Raw and average spectra of TZS samples in the range of Vis/NIR and SWIR. **(A)** Raw spectra and **(C)** average spectra of TZS samples in the range of Vis/NIR; **(B)** Raw spectra and **(D)** average spectra of TZS samples in the range of SWIR.

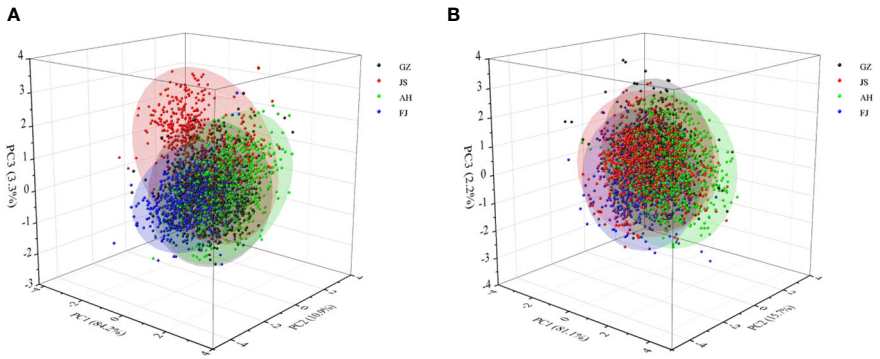


FIGURE 3
Scores scatter plots of Vis/NIR and SWIR spectra of TZS from four geographical origins. **(A)** Vis/NIR spectra; **(B)** SWIR spectra.

on the spectral data from the Vis/NIR and SWIR regions as the input matrices, respectively (Table 2). The loss and accuracy curves for the optimal model were depicted in Figure 4. The loss value of the discriminative model continuously decreased as the number of iterations increased. Still, the precision increased and ultimately stabilized, indicating that the FD-DCNN model converged properly. In all cases, the discriminative performance of the models using SWIR spectra was superior to those using Vis/NIR spectra. The prediction accuracies of the models in the Vis/NIR region were lower than 91.4%, while the SWIR models could reach 98.7% accuracies. The selection of sensor type (information source), pretreatment methods, and classifier collectively influence the discrimination accuracy of the models to varying extents. Compared to the SNV and DT algorithms, the FD algorithm exhibited superior preprocessing performance in both the visible and near-infrared regions. This suggested that FD might be an ideal preprocessing method to improve the signal-to-noise ratio of spectra associated with the origin of TZS. Additionally, in the

TABLE 2 Results of classification models based on individual spectral region datasets.

Ranges	Models	Treatments	Parameters	Classification accuracy (%)	
				Training set	Validation set
Vis/NIR	PLS-DA	Raw	LV: 12	88.6	86.1
		SNV	LV: 9	84.3	82.9
		DT	LV: 6	66.6	66.7
		FD	LV: 10	88.3	85.7
	RF	Raw	T: 200; L: 1	100.0	71.1
		SNV	T: 200; L: 1	100.0	81.6
		DT	T: 200; L: 1	100.0	79.1
		FD	T: 200; L: 1	100.0	85.1
	SVM	Raw	C: 1000000.0, gamma: 0.0001	98.0	56.2
		SNV	C: 100000.0, gamma: 0.0001	98.1	80.5
		DT	C: 1000000.0, gamma: $1e^{-05}$	98.0	73.7
		FD	C: 100000.0, gamma: $1e^{-06}$	96.1	90.0
	DCNN	Raw	—	68.9	68.5
		SNV	—	64.7	63.4
		DT	—	70.5	69.4
		FD	—	94.8	91.4

(Continued)

10.3389/fpls.2023.1342970

TABLE 2 Continued

Ranges	Models	Treatments	Parameters	Classification accuracy (%)	
				Training set	Validation set
SWIR	PLS-DA	Raw	LV: 7	86.6	87.8
		SNV	LV: 8	88.6	90.9
		DT	LV: 8	88.7	90.8
		FD	LV: 8	88.2	88.7
	RF	Raw	T: 200; L: 1	100.0	79.4
		SNV	T: 200; L: 1	100.0	91.5
		DT	T: 200; L: 1	100.0	90.9
		FD	T: 200; L: 1	100.0	96.8
	SVM	Raw	C: 464158.8, gamma: $1e^{-05}$	97.1	76.1
		SNV	C: 100000.0, gamma: 0.0001	98.9	68.8
		DT	C: 10000000.0, gamma: $1e^{-06}$	98.3	95.0
		FD	C: 10000.0, gamma: $1e^{-05}$	98.3	97.6
	DCNN	Raw	—	68.3	67.6
		SNV	—	93.0	93.4
		DT	—	94.9	94.1
		FD	—	99.6	98.7

SWIR range, the model built with DCNN combined with FD pretreatment algorithm exhibited the highest precision, achieving 98.7% discrimination accuracy for the four origins of the TZS samples. Interestingly, the FD-SVM and FD-RF models also obtained satisfactory classification accuracies with validation sets of 97.6% and 96.8%, respectively. The confusion matrices of the models for the SWIR region were illustrated in Figure 5, which revealed that the FD-SVM and FD-DCNN models not only achieved desirable accuracies in terms of total correctness, but their accuracies were still high (>95.6%) for each origin category. With regard to the application of the model, we sought to reduce both the associated equipment costs and the time required for model prediction. Consequently, further research was carried out to extract a smaller number of feature bands from the SWIR spectra to establish a more efficient discrimination method.

3.4 Selection of effective wavelengths

An appropriate wavelength selection method is crucial as it not only reduces the number of wavebands but also helps eliminate irrelevant, noisy, or collinear variables, thereby improving the modeling precision (Liu et al., 2022). Moreover, different wavelength selection methods are based on different algorithm principles, which can lead to varying modeling results when applied to different types of datasets. It is important to carefully consider the characteristics of the dataset and choose a wavelength selection method that best suits the specific needs. In this study, CARS and SPA were utilized to select the effective wavelengths (EWs) from the SWIR spectra that would potentially contain the most valuable information regarding the geographical origins of TZS samples.

The randomness of the Monte Carlo sampling resulted in inconsistent results for every operation of the CARS approach and the optimal results after 10 CARS selections were chosen to obtain a relatively optimal combination of bands (Figure 6). Under the exponential decay function, the number of bands decreases rapidly at the beginning of the sampling, but with the sampling number increasing, the rate of decrease of the band number slows down gradually (Figure 6A). As shown in Figure 6B, the RMSECV values showed an overall trend of decreasing and then increasing with the sampling times, and the RMSECV values were the lowest when the number of sampling times reached 47. Combined with Figure 6C, it was observed that the RMSECV value was the smallest at the 47th sampling (* denotes), meaning that the subset containing 32 variables selected for this sampling was the key to determining the origins of TZS. The SPA method establishes a multiple linear regression model for different subsets of bands one by one and calculates the RMSEP values when selecting the optimal bands, where the subset corresponding to the smallest value is the optimal subset of bands. As shown in Figure 6D, the RMSEP values show an overall decreasing trend with the increase in the number of

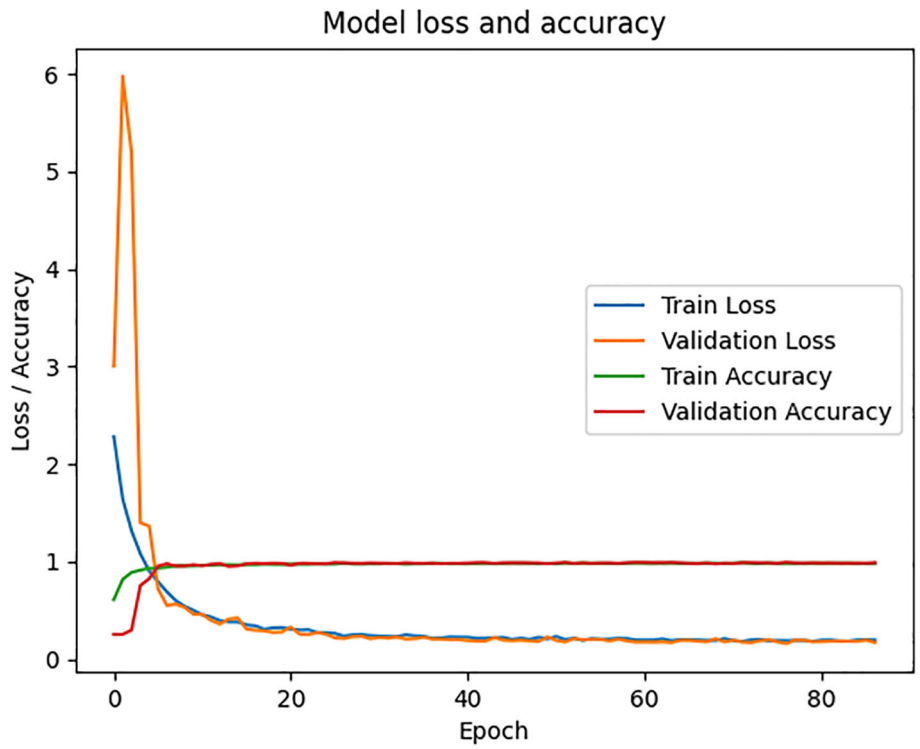


FIGURE 4
The loss and accuracy curves of the FD-DCNN model based on the SWIR.

bands. When the number of bands reaches 110, the RMSEP value minimizes to 0.444 and then slightly increases. The specific descriptions of the EWs screened with the CARS and SPA algorithms are listed in Table 3.

3.5 Classification models based on EWs

After applying the CARS and SPA algorithms to select the essential wavelengths (EWs), simplified PLS-DA, RF and SVM models were developed to determine the geographic origins of TZS (Table 4 and Figure 7). The models exhibited different performances, indicating that the choice of wavelength selection method had varying effects on the discriminative models of TZS origins. The performance of the CARS-RF and SPA-RF models exhibited a slight degradation compared to the full-band FD-RF model. Additionally, both SVM and RF models based on the CARS method performed worse than the SPA-based SVM and RF models. This discrepancy might be attributed to the limited number of EWs selected by CARS, leading to the elimination of some EWs containing crucial information about the TZS origins. Remarkably, the SPA-SVM model based on 110 EWs obtained the optimal discriminative accuracy with 98.4% for the training set and 98.1% for the validation set. Although it was slightly worse than the full-band FD-DCNN model (training set of 99.6% and validation set of 98.7%), it outperformed the full-band FD-SVM model (training set of 98.3% and validation set of 97.6%). These results indicated that compared to CARS, the SPA algorithm is preferable for extracting the SWIR information that is highly correlated with the TZS origins.

An additional analysis was performed on the extracted EWs from SPA, as shown in Table 3. This analysis revealed that most of the EWs were concentrated in specific regions of the spectra, indicating a potential relationship between the origin and chemical composition of TZS. The wavebands around 970 nm are associated with O-H second overtone stretching vibration and C-H stretching third overtone, which are related to sugar and cellulose, respectively (Theanjumpol et al., 2013). The bands between 1050 nm and 1200 nm, as well as 1300 nm and 1500 nm, are the main characteristic spectral regions that represent the 20 amino acids found in proteins. The 1050-1200 nm region primarily consists of the second overtone of C-H, while the 1300-1500 nm region is mainly composed of the combined frequency of C-H, reflecting the differences in amino acid composition among different samples (Weinstock et al., 2006; Jin et al., 2022).

3.6 Optimal model validation and visualization

Apart from the 3249 TZS samples used for modeling, an additional 320 samples (80 TZS per origin) were selected for external validation and visualization of the optimal model (FD-DCNN). The visualization of the validation results is shown in Figure 8. The origin of TZS was marked with different colors, with red representing TZS identified by the model as originating from Guizhou (GZ) Province, pink

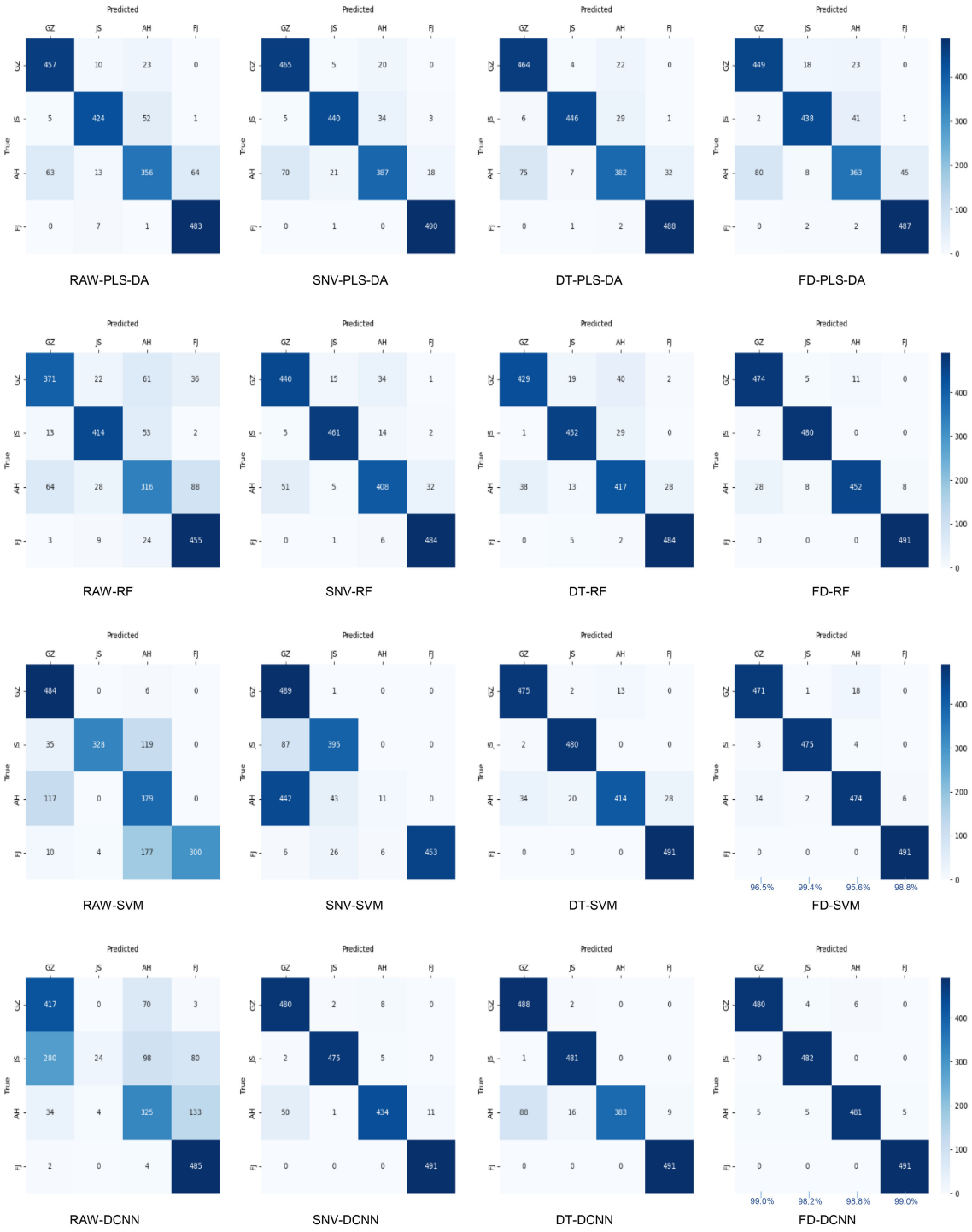


FIGURE 5
The confusion matrices of the PLS-DA, SVM, RF and DCNN models on the prediction set using different preprocessed SWIR spectra.

representing TZS identified by the model as originating from Jiangsu (JS) Province, purple representing TZS identified as originating from Anhui (AH) Province, and blue indicating TZS identified as originating from Fujian (FJ) Province. It can be seen that all TZS from Guizhou, Jiangsu, and Fujian provinces were correctly recognized (100%). One sample in the Anhui group was incorrectly identified as TZS from Guizhou Province with a precision of 98.8%. The results of this external validation were consistent with the results of the FD-DCNN model, indicating that the discrimination model developed in this study for TZS had excellent robustness.

Previous studies on the discrimination of the origin of TZS are based on only one variety from one origin (Wu et al., 2018; Pan et al., 2020), overlooking the disturbances caused by the genetic background and grade differences, which leads to the

10.3389/fpls.2023.1342970

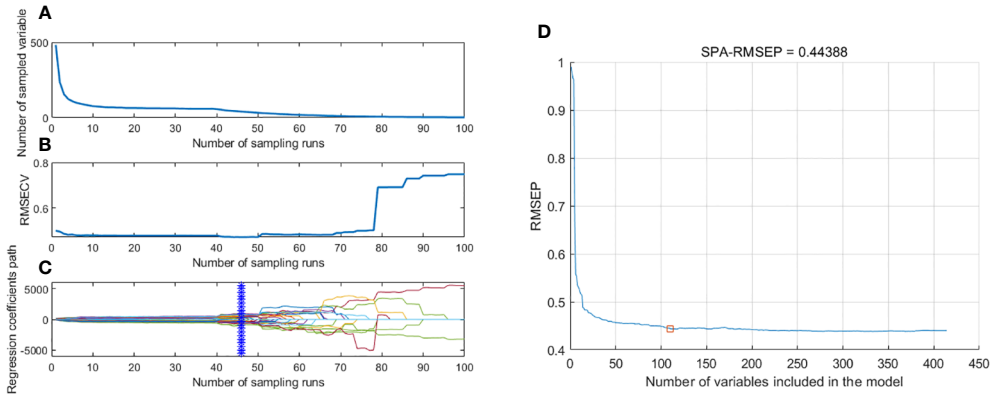


FIGURE 6
The process of extracting EWs with CARS and SPA. **(A)** Number of preferred EWs with CARS; **(B)** The root mean square error of cross-validation variation with CARS; **(C)** Regression coefficient path map with CARS; **(D)** Extraction of EWs with SPA.

limited application of the model. Our methodology considered the representativeness of the samples and the applicability of the approach by selecting two or three germplasm resources for each geographical region. Furthermore, each germplasm resource included different quality grades of TZS (Table 1), enhancing the comprehensiveness of the analysis.

In the analysis of the two spectral ranges used in this study, the models within the range of 900.96~1700.43 nm demonstrated superior performance compared to the models within the 400.20~999.75 nm range. The correct classification rates for the prediction set ranged from 67.6% to 98.1% for the 900.96~1700.43 nm range (Tables 2, 4), while they ranged from 56.2% to 91.4% for the 400.20~999.75 nm range (Table 2). This difference in accuracy can be attributed to the fact that the spectra in the 900.96~1700.43 nm range provide information about the stretching vibrations of C-H, O-H, and N-H, which are caused by starch, protein, cellulose, and water in the TZS. On the other hand, the

TABLE 3 Specific description of the selected EWs by SPA and CARS.

Methods	No.	EWs							
CARS	32	919.24	930.88	934.2	935.86	959.13	969.11	975.75	1117.03
		1130.33	1131.99	1143.63	1148.61	1150.28	1168.56	1175.21	1245.02
		1286.57	1341.42	1377.98	1381.31	1382.97	1384.63	1391.28	1407.9
		1409.56	1417.88	1421.2	1434.5	1574.11	1632.29	1635.61	1640.6
SPA	110	907.61	909.27	910.93	922.57	924.23	929.21	934.2	940.85
		945.84	947.5	949.16	954.15	955.81	959.13	965.78	969.11
		987.39	989.05	990.71	1002.35	1004.01	1005.67	1007.33	1012.32
		1017.31	1018.97	1020.63	1025.62	1027.28	1032.27	1033.93	1035.59
		1043.9	1048.89	1050.55	1058.86	1062.18	1063.85	1065.51	1078.8
		1080.47	1092.1	1093.76	1095.43	1105.4	1108.72	1110.38	1113.71
		1115.37	1118.7	1120.36	1128.67	1131.99	1138.64	1145.29	1151.94
		1158.59	1188.5	1213.44	1215.1	1218.42	1228.39	1230.06	1240.03
		1241.69	1245.02	1256.65	1283.24	1284.91	1311.5	1313.16	1328.12
		1329.78	1333.11	1341.42	1346.4	1349.73	1361.36	1364.69	1366.35
		1369.67	1376.32	1377.98	1384.63	1391.28	1407.9	1414.55	1419.54
		1431.17	1432.83	1436.16	1437.82	1446.13	1487.68	1489.35	1491.01
		1492.67	1495.99	1502.64	1504.3	1505.97	1509.29	1510.95	1584.09
		1600.71	1628.96	1638.94	1643.92	1645.58	1648.91		

TABLE 4 Results of simplified classification models based on SWIR spectra.

Models	EWs selection methods	Number of EWs	Parameters	Classification accuracy (%)	
				Training set	Validation set
PLS-DA	CARS	32	LV: 7	86.1	87.9
	SPA	110	LV: 9	89.2	90.9
RF	CARS	32	T: 50; L: 1	100.0	95.8
	SPA	110	T: 50; L: 1	99.9	97.0
SVM	CARS	32	C:10000000.0, gamma: $1e^{-06}$	97.6	84.7
	SPA	110	C:10000.0, gamma: 0.0001	98.4	98.1

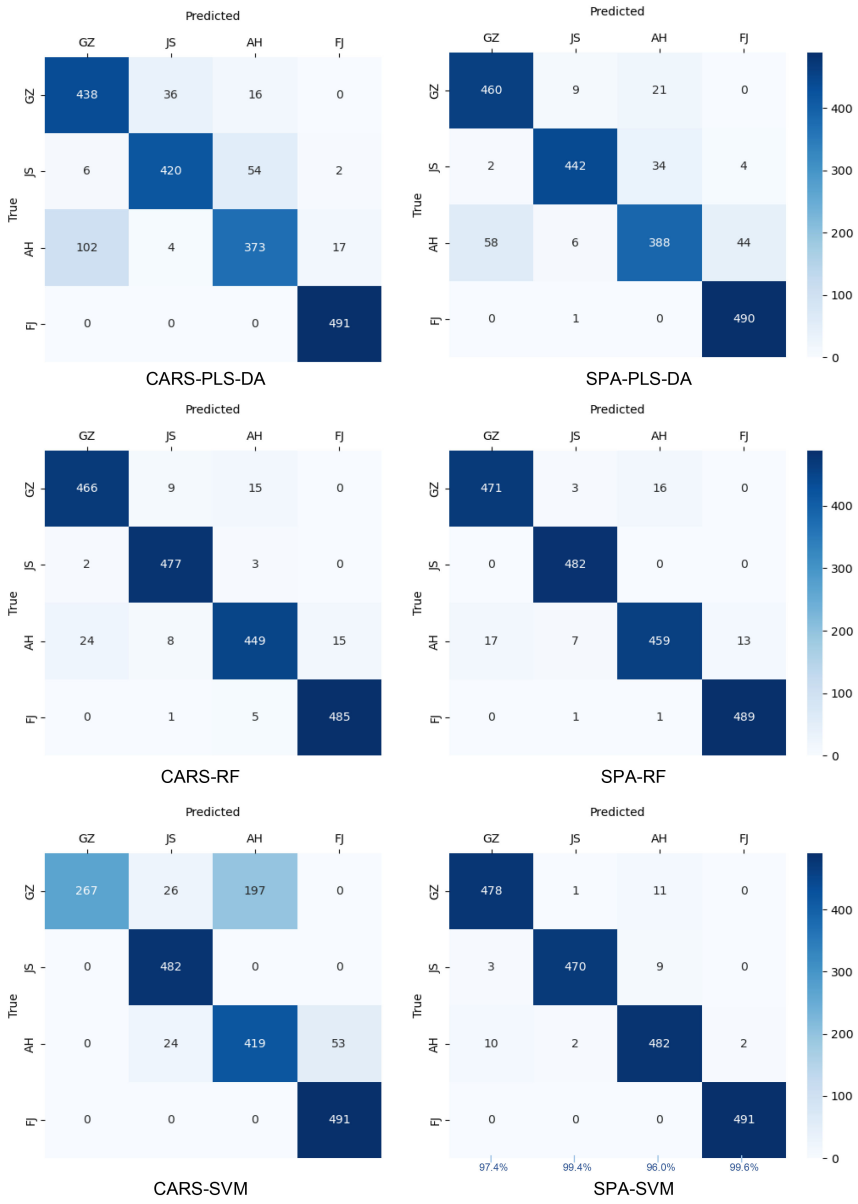


FIGURE 7
The confusion matrices of the simplified PLS-DA, SVM and RF models on the prediction set.

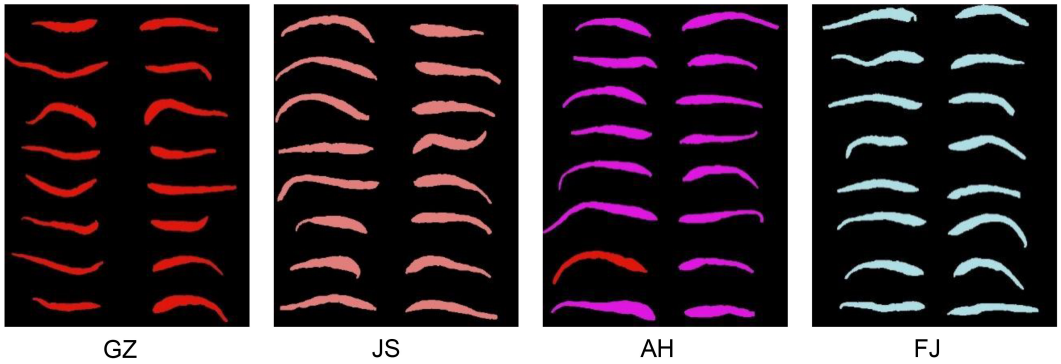


FIGURE 8
Detection visualization of TZS samples from Guizhou (GZ), Jiangsu (JS), Anhui (AH) and Fujian (FJ) Provinces.

wavelengths between 400.20 and 999.75 nm primarily reflect the color and pigment information in the TZS. Hou et al. (2015) conducted an analysis of the chemical compositions in P. *heterophylla* from different origins using UPLC-Triple TOF-MS/MS. The study identified 21 distinct chemical components, including maltotriose, sucrose, thyronine, inosine triphosphate, pseudostellarin A, pseudostellarin B, pseudostellarin D, pseudostellarin F, heterophyllin A and sphinganine. Compared with other origins, the levels of pseudostellarin D, pseudostellarin E, pesudostellarin A, heterophyllin A, pseudostellarin F, isobutyrylglycine in P. *heterophylla* from Fujian were higher. Sucrose, ferulic acid, canthaxanthin, maltotriose, pseudostellarin D in P. *heterophylla* from Guizhou were richer than those of other origins (Sha et al., 2023). Hence, it is reasonable to hypothesize that spectral differences resulting from variations in chemical composition, rather than color and pigmentation information, may play a crucial role in studying the traceability of the origin of *Pseudostellaria heterophylla*.

Notably, our work takes a novel approach by using hyperspectral imaging (HSI) in conjunction with deep learning (DL) techniques to assess the geographical origins of TZS. Wu et al. (2018) highlighted the efficiency of using Raman spectroscopy combined with MSC-SG-CARS-PLS-DA to discriminate P. *heterophylla* from different regions. Similarly, Pan et al. (2020) demonstrated that NIR spectroscopy in combination with Row-center-SG-CARS-PLS-DA could be effective in distinguishing the P. *heterophylla* from different regions. Further to this, this research conducted a comparison between two feature band extraction algorithms, namely CARS and SPA. The results showed that the SPA algorithm was preferable for extracting SWIR information, which was highly correlated with the TZS origins (Table 3 and Table 4). Furthermore, we compared the traditional two-stage machine learning algorithms (PLS-DA, SVM, and RF) with the end-to-end deep learning algorithm (DCNN). Our findings demonstrated that both SVM and DCNN classifiers outperformed PLS-DA and RF classifiers in terms of origin identification of TZS (Table 2 and Figure 5). Several previous studies indicated that nonlinear models, such as SVM, were superior to linear models in solving the seed classifications (Qiu et al., 2018; Wakholi et al., 2018; Zhao et al., 2018). For the first time, our work further argued this on geographic origin recognition in TZS.

4 Conclusion

In this study, the visible near-infrared (Vis/NIR) and short-wave infrared (SWIR) hyperspectral information from different origins of TZS samples were collected. By combining various preprocessing algorithms, feature band extraction algorithms, traditional two-stage machine learning, and end-to-end deep learning classifiers, we developed fast and high-precision identification methods to discriminate TZS origins. The specific conclusions drawn from this study are as follows:

1) The model accuracy based on SWIR HSI for identifying the geographical origins of TZS was higher compared to that based on Vis/NIR HSI. The best model accuracy using Vis/NIR HSI was 91.4%, while the optimal model accuracy using SWIR HSI could reach up to 98.7%.
2) The SPA algorithm was suitable for extracting SWIR information, which was highly correlated with the origins of TZS. The corresponding FD-SPA-SVM model not only reduced the number of bands by 77.2% but also improved the model accuracy from 97.6% to 98.1% compared to the full-band FD-SVM model.
3) Two sets of fast and high-precision methods were developed to distinguish between different geographic origins of TZS. The traditional two-stage machine learning classifier achieves optimal performance by employing the SVM model with FD pretreatment and the variable selection method of SPA. In contrast, the end-to-end deep learning classifier achieves optimal discrimination by solely applying FD preprocessing combined with DCNN. The total accuracies of the SWIR-FD-SPA-SVM model and the

SWIR-FD-DCNN model for identifying TZS origins were 98.1% and 98.7%, respectively.

This work provides a potentially perfect tool for herbal companies and market regulators to widely identify the origins of TZS across various genetic backgrounds and quality grades.

Data availability statement

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

Author contributions

TZ: Conceptualization, Methodology, Supervision, Visualization, Writing – original draft, Writing – review & editing. LL: Data curation, Formal analysis, Writing – review & editing. YS: Data curation, Writing – original draft. MY: Writing – review & editing. JL: Writing – review & editing. JY: Writing – review & editing. YL: Writing – review & editing. XS: Writing – review & editing. ML: Writing – review & editing. XY: Writing – review & editing. ZZ: Writing – review & editing. RZ: Writing – review & editing. YuS: Funding acquisition, Supervision, Writing – review & editing. LG: Funding acquisition, Supervision, Writing – review & editing.

Funding

The author(s) declare financial support was received for the research, authorship, and/or publication of this article. This research was supported by the National Natural Science Foundation of China (82373994 and 32371588), the China agriculture research system of MOF and MARA (CARS-21), the Natural Science Foundation of Fujian Province, China (2021J02024).

Conflict of interest

Authors YL and XS were employed by the company Huzhou Wuxing Jinnong Ecological Agriculture Development 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

References

Alchanatis, V., Schmilovitch, Z., and Meron, M. (2005). In-field assessment of single leaf nitrogen status by spectral reflectance measurements. *Precis. Agric.* 6, 25–39. doi: 10.1007/s11119-005-0682-7

Barnes, R., Dhanoa, M. S., and Lister, S. J. (1989). Standard normal variate transformation and de-trending of near-infrared diffuse reflectance spectra. *Appl. Spectrosc.* 43, 772–777. doi: 10.1366/0003702894202201

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

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

Delwiche, S. R., and Kim, M. S. (2000). Hyperspectral imaging for detection of scab in wheat. *Biol. Qual. Precis. Agric. II.* 4203, 13–20. doi: 10.1117/12.411752

Dong, F., Hao, J., Luo, R., Zhang, Z., Wang, S., Wu, K., et al. (2022). Identification of the proximate geographical origin of wolfberries by two-dimensional correlation spectroscopy combined with deep learning. *Comput. Electron. Agric.* 198, 107027. doi: 10.1016/j.compag.2022.107027

Fu, L., Feng, Y., Majeed, Y., Zhang, X., Zhang, J., Karkee, M., et al. (2018). Kiwifruit detection in field images using Faster R-CNN with ZFNet. *IFAC-PapersOnLine* 51, 45–50. doi: 10.1016/j.ifacol.2018.08.059

Hou, Y., Ma, Y., Zou, L., Liu, X., Liu, X., Luo, Y., et al. (2015). Difference of chemical compositions in Pseudostellariae Radix from different origins by UPLC-Triple TOF–MS/MS. *J. Chin. Mass Spectrom. Soc* 36, 359–366. doi: 10.7538/zpxb.youxian.2015.0019

Ioffe, S., and Szegedy, C. (2015). Batch normalization: Accelerating deep network training by reducing internal covariate shift. *Int. Conf. Mach. Learn.* 37, 448–456. doi: 10.48550/arXiv.1502.03167

Jin, B., Zhang, C., Jia, L., Tang, Q., Gao, L., Zhao, G., et al. (2022). Identification of rice seed varieties based on near-infrared hyperspectral imaging technology combined with deep learning. *ACS omega.* 7 (6), 4735–4749. doi: 10.1021/acsomega.1c04102

Kumar, T. B., Prashar, D., Vaidya, G., Kumar, V., Kumar, S., and Sammy, F. (2022). A novel model to detect and classify fresh and damaged fruits to reduce food waste using a deep learning technique. *J. Food Qual.* 2022. doi: 10.1155/2022/4661108

Li, J., Zhen, W., Long, D., Ding, L., Gong, A., Xiao, C., et al. (2016). *De novo* sequencing and assembly analysis of the *Pseudostellaria heterophylla* transcriptome. *PLoS One* 11, e0164235. doi: 10.1371/journal.pone.0164235

Lin, H., Zhao, J., Chen, Q., Zhou, F., and Sun, L. (2011). Discrimination of Radix *Pseudostellariae* according to geographical origins using NIR spectroscopy and support vector data description. *Spectrochim. Acta A Mol. Biomol. Spectrosc.* 79, 1381–1385. doi: 10.1016/j.saa.2011.04.072

Liu, Z., Jiang, J., Qiao, X., Qi, X., Pan, Y., and Pan, X. (2020). Using convolution neural network and hyperspectral image to identify moldy peanut kernels. *LWT* 132, 109815. doi: 10.1016/j.lwt.2020.109815

Liu, C., Wang, Q., Lin, W., and Yu, C. (2022). Origins classification of egg with different storage durations using FT-NIR: A characteristic wavelength selection approach based on information entropy. *Biosyst. Eng.* 222, 82–92. doi: 10.1016/j.biosystemseng.2022.07.016

López-Maestresalas, A., Keresztes, J. C., Goodarzi, M., Arazuri, S., Jarén, C., and Saeys, W. (2016). Non-destructive detection of blackspot in potatoes by Vis-NIR and SWIR hyperspectral imaging. *Food Control.* 70, 229–241. doi: 10.1016/j.foodcont.2016.06.001

Lu, Y., Saeys, W., Kim, M., Peng, Y., and Lu, R. (2020). Hyperspectral imaging technology for quality and safety evaluation of horticultural products: A review and celebration of the past 20-year progress. *Postharvest Biol. Technol.* 170, 111318. doi: 10.1016/j.postharvbio.2020.111318

Nie, P., Zhang, J., Feng, X., Yu, C., and He, Y. (2019). Classification of hybrid seeds using near-infrared hyperspectral imaging technology combined with deep learning. *Sens. Actuators B Chem.* 296, 126630. doi: 10.1016/j.snb.2019.126630

Pan, W., Wu, M., Zheng, Z., Guo, L., Lin, Z., and Qiu, B. (2020). Rapid authentication of *Pseudostellaria heterophylla* (Taizishen) from different regions by near-infrared spectroscopy combined with chemometric methods. *J. Food Sci.* 85, 2004–2009. doi: 10.1111/1750-3841.15171

Qiu, Z., Chen, J., Zhao, Y., Zhu, S., He, Y., and Zhang, C. (2018). Variety identification of single rice seed using hyperspectral imaging combined with convolutional neural network. *Appl. Sci.* 8, 212. doi: 10.3390/app8020212

Sha, M., Li, X., Liu, Y., Tian, H., Liang, X., Li, X., et al. (2023). Comparative chemical characters of *Pseudostellaria heterophylla* from geographical origins of China. *Chin. Herb. Med.* 15 (3), 439–446. doi: 10.1016/j.chmed.2022.10.005

Shi, X., Jian, W., Wang, Z., and Li, H. (2013). Overview of cultivation technology of *Pseudostellaria heterophylla*. Strait. *Pharm. J.* 25, 15–18.

Theanjumpol, P., Self, G., Rittiron, R., Pankasemsuk, T., and Sardsud, V. (2013). Selecting variables for near infrared spectroscopy (NIRS) evaluation of mango fruit quality. *J. Agric. Sci.* 5 (7), 146. doi: 10.5539/jas.v5n7p146

Tian, X., Liu, X., He, X., Zhang, C., Li, J., and Huang, W. (2023). Detection of early bruises on apples using hyperspectral reflectance imaging coupled with optimal wavelengths selection and improved watershed segmentation algorithm. *J. Sci. Food Agric.* 103 (13), 6689–6705. doi: 10.1002/jsfa.12764

Tian, X., Zhang, C., Li, J., Fan, S., Yang, Y., and Huang, W. (2021). Detection of early decay on citrus using LW-NIR hyperspectral reflectance imaging coupled with two-band ratio and improved watershed segmentation algorithm. *Food Chem.* 360, 130077. doi: 10.1016/j.foodchem.2021.130077

Wakholi, C., Kandpal, L. M., Lee, H., Bae, H., Park, E., Kim, M. S., et al. (2018). Rapid assessment of corn seed viability using short wave infrared line-scan hyperspectral imaging and chemometrics. *Sens. Actuators B Chem.* 255, 498–507. doi: 10.1016/j.snb.2017.08.036

Wang, Z., Huang, W., Li, J., Liu, S., and Fan, S. (2023). Assessment of protein content and insect infestation of maize seeds based on on-line near-infrared spectroscopy and machine learning. *Comput. Electron. Agric.* 211, 107969. doi: 10.1016/j.compag.2023.107969

Wang, Z., Huang, W., Tian, X., Long, Y., Li, L., and Fan, S. (2022). Rapid and non-destructive classification of new and aged maize seeds using hyperspectral image and chemometric methods. *Front. Plant Sci.* 13, 849495. doi: 10.3389/fpls.2022.849495

Weinstock, B. A., Janni, J., Hagen, L., and Wright, S. (2006). Prediction of oil and oleic acid concentrations in individual corn (*Zea mays* L.) kernels using near-infrared reflectance hyperspectral imaging and multivariate analysis. *Appl. Spectrosc.* 60 (1), 9–16. doi: 10.1366/000370206775382631

Wong, C., Leung, K., Fung, K., and Choy, Y. (1994). The immunostimulating activities of anti-tumor polysaccharides from *Pseudostellaria heterophylla*. *Immunopharmacology* 28, 47–54. doi: 10.1016/0162-3109(94)90038-8

Wu, M., Chen, L., Huang, X., Zheng, Z., Qiu, B., Guo, L., et al. (2018). Rapid authentication of *Pseudostellaria heterophylla* (Taizishen) from different regions by Raman spectroscopy coupled with chemometric methods. *J. Lumin.* 202, 239–245. doi: 10.1016/j.jlumin.2018.05.036

Yu, Z., Fang, H., Zhang, Q., Mi, C., Feng, X., and He, Y. (2021). Hyperspectral imaging technology combined with deep learning for hybrid okra seed identification. *Biosyst. Eng.* 212, 46–61. doi: 10.1016/j.biosystemseng.2021.09.010

Zareef, M., Arslan, M., Hassan, M. M., Ahmad, W., Ali, S., Li, H., et al. (2021). Recent advances in assessing qualitative and quantitative aspects of cereals using nondestructive techniques: A review. *Trends Food Sci. Technol.* 116, 815–828. doi: 10.1016/j.tifs.2021.08.012

Zhang, T., Fan, S., Xiang, Y., Zhang, S., Wang, J., and Sun, Q. (2020b). Non-destructive analysis of germination percentage, germination energy and simple vigour index on wheat seeds during storage by Vis/NIR and SWIR hyperspectral imaging. *Spectrochim. Acta A Mol. Biomol. Spectrosc.* 239, 118488. doi: 10.1016/j.saa.2020.118488

Zhang, H., Hou, Q., Luo, B., Tu, K., Zhao, C., and Sun, Q. (2022b). Detection of seed purity of hybrid wheat using reflectance and transmittance hyperspectral imaging technology. *Front. Plant Sci.* 13, 1015891. doi: 10.3389/fpls.2022.1015891

Zhang, T., Wei, W., Zhao, B., Wang, R., Li, M., Yang, L., et al. (2018). A reliable methodology for determining seed viability by using hyperspectral data from two sides of wheat seeds. *Sensors* 18, 813. doi: 10.3390/s18030813

Zhang, C., Wu, W., Zhou, L., Cheng, H., Ye, X., and He, Y. (2020a). Developing deep learning based regression approaches for determination of chemical compositions in dry black goji berries (*Lycium ruthenicum* Murr.) using near-infrared hyperspectral imaging. *Food Chem.* 319, 126536. doi: 10.1016/j.foodchem.2020.126536

Zhang, Q., Zhang, M., Chen, T., Sun, Z., Ma, Y., and Yu, B. (2019). Recent advances in convolutional neural network acceleration. *Neurocomputing* 323, 37–51. doi: 10.1016/j.neucom.2018.09.038

Zhang, C., Zhou, L., Xiao, Q., Bai, X., Wu, B., Wu, N., et al. (2022a). End-to-end fusion of hyperspectral and chlorophyll fluorescence imaging to identify rice stresses. *Plant Phenomics.* 2022. doi: 10.34133/2022/9851096

Zhao, Y., Zhu, S., Zhang, C., Feng, X., Feng, L., and He, Y. (2018). Application of hyperspectral imaging and chemometrics for variety classification of maize seeds. *Rsc Advances.* 8, 1337–1345. doi: 10.1039/C7RA05954J

frontiers | Frontiers in Plant Science

TYPE Original Research
PUBLISHED 23 January 2024
DOI 10.3389/fpls.2024.1324753

OPEN ACCESS

EDITED BY
Te Ma,
Nagoya University, Japan

REVIEWED BY
Baohua Zhang,
Nanjing Agricultural University, China
Zhang Hailiang,
East China Jiaotong University, China
Rajko Vidrih,
University of Ljubljana, Slovenia
Yi Yang,
Beijing Technology and Business University, China
Leiqing Pan,
Nanjing Agricultural University, China

*CORRESPONDENCE
Xuhai Yang
yxh_513@shzu.edu.cn
Qian Zhang
zq_mac@shzu.edu.cn

†These authors have contributed equally to this work

RECEIVED 20 October 2023
ACCEPTED 08 January 2024
PUBLISHED 23 January 2024

CITATION
Li S, Li J, Wang Q, Shi R, Yang X and Zhang Q (2024) Determination of soluble solids content of multiple varieties of tomatoes by full transmission visible-near infrared spectroscopy.
Front. Plant Sci. 15:1324753.
doi: 10.3389/fpls.2024.1324753

Determination of soluble solids content of multiple varieties of tomatoes by full transmission visible-near infrared spectroscopy

Sheng Li[1,2,3,4,5†], Jiangbo Li[2†], Qingyan Wang[2], Ruiyao Shi[2], Xuhai Yang[1,3,4,5*] and Qian Zhang[1,3,4,5*]

[1]College of Mechanical and Electrical Engineering, Shihezi University, Shihezi, China, [2]Intelligent Equipment Research Center, Beijing Academy of Agriculture and Forestry Sciences, Beijing, China, [3]Xinjiang Production and Construction Corps Key Laboratory of Modern Agricultural Machinery, Shihezi, China, [4]Key Laboratory of Northwest Agricultural Equipment, Ministry of Agriculture and Rural Affairs, Shihezi, China, [5]Engineering Research Center for Production Mechanization of Oasis Characteristic Cash Crop, Ministry of Education, Shihezi, China

Introduction: Soluble solids content (SSC) is a pivotal parameter for assessing tomato quality. Traditional measurement methods are both destructive and time-consuming.

Methods: To enhance accuracy and efficiency in SSC assessment, this study employs full transmission visible and near-infrared (Vis-NIR) spectroscopy and multi-point spectral data collection techniques to quantitatively analyze SSC in two tomato varieties ('Provence' and 'Jingcai No.8' tomatoes). Preprocessing of the multi-point spectra is carried out using a weighted averaging approach, aimed at noise reduction, signal-to-noise ratio improvement, and overall data quality enhancement. Taking into account the potential influence of various detection orientations and preprocessing methods on model outcomes, we investigate the combination of partial least squares regression (PLSR) with two orientations (O1 and O2) and two preprocessing techniques (Savitzky-Golay smoothing (SG) and Standard Normal Variate transformation (SNV)) in the development of SSC prediction models.

Results: The model achieved the best results in the O2 orientation and SNV pretreatment as follows: 'Provence' tomato (Rp = 0.81, RMSEP = 0.69°Brix) and 'Jingcai No.8' tomatoes (Rp = 0.84, RMSEP = 0.64°Brix). To further optimize the model, characteristic wavelength selection is introduced through Least Angle Regression (LARS) with L1 and L2 regularization. Notably, when λ=0.004, LARS-L1 produces superior results ('Provence' tomato: Rp = 0.95, RMSEP = 0.35°Brix; 'Jingcai No.8' tomato: Rp = 0.96, RMSEP = 0.33°Brix).

Discussion: This study underscores the effectiveness of full transmission Vis-NIR spectroscopy in predicting SSC in different tomato varieties, offering a viable method for accurate and swift SSC assessment in tomatoes.

KEYWORDS

tomato, soluble solids content, online detection, full transmission, quantitative analysis model

1 Introduction

Tomatoes, ubiquitous in global agricultural production, exhibit noteworthy nutritional significance (Passam et al., 2007). As an esteemed vegetable within the realm of human dietary practices, tomatoes assume a pivotal role (Zhu et al., 2015). They bestow a diverse complement of indispensable organic compounds, thus exerting a multifaceted influence encompassing pigmentation modulation, retardation of aging processes, lipid and blood pressure reduction, safeguarding of prostatic health, as well as reinforcement of gastric and hepatic functions (Perveen et al., 2015; Youssef and Eissa, 2017; Salehi et al., 2019). By virtue of their unique attributes, encompassing edibility, health-enhancing properties, and therapeutic potential, tomatoes and their derivative products occupy a prominent and indispensable position within the global landscape of agricultural production and trade (Guan et al., 2018; Ali et al., 2020). Soluble solids content (SSC) represents a pivotal constituent of tomato flavor, holding the potential to align closely with consumers' perception of intrinsic quality attributes in tomatoes (Ponce-Valadez et al., 2016). Nevertheless, the conventional analytical methodologies employed for quantifying this quality parameter are beset with challenges related to protracted analysis durations, substantial costs, and environmental contamination (Skolik et al., 2019). The imperative for the tomato production and distribution industry, therefore, resides in the development of expeditious, facile, cost-effective, environmentally benign, and non-invasive techniques for batch quality control assessment, with the ability to extend precision down to the level of individual fruits (Najjar and Abu-Khalaf, 2021).

In the past few decades, many non-destructive testing techniques have been used to detect tomato SSC (Mei and Li, 2023). Gómez et al. (2008) used PEN 2 electronic nose (E-nose) to detect tomatoes with different storage time. The correlation between the measured value and the predicted value showed that the effect of using E-nose sensor signal to predict tomato SSC was poor. Nikbakht et al. (2011) used raman spectroscopy to determine the SSC of tomato. The root mean square error of predictions (RMSEP) of SSC measured by partial least squares regression (PLSR) and principal component regression (PCR) models were 0.30 and 0.38, respectively. In order to explore the possibility of mid-infrared spectroscopy for tomato quality detection, Ścibisz et al. (2011) used the attenuated total reflection accessory of the fourier transform spectrometer to scan the tomato samples in the wavenumber region of 4000 to 400 cm^{-1}. The PLSR model has a reasonable ability to estimate the SSC of tomatoes, with a high determination coefficient (0.98) and a small prediction error (3%). Mollazade et al. (2015) used backscattering and multispectral imaging techniques to predict the quality factors of tomato fruit during storage. The correlation coefficients between the prediction results of SSC correction model established by artificial neural network and the reference measurement results of multispectral and backscatter imaging are 0.736 and 0.561, respectively. Rahman et al. (2017) established a non-destructive method for the determination of SSC in intact tomatoes using hyperspectral imaging technology in the range of 1000-1550 nm. The PLSR model based on smoothing pretreatment spectrum has a good prediction effect on SSC of intact tomatoes, with a correlation coefficient of prediction (Rp) of 0.74 and a RMSEP of 0.33%.

While the aforementioned methods enable non-destructive testing, their inherent time-consuming nature and elevated cost factors constrain their utility when catering to the rigorous industrial testing requisites characteristic of large-scale tomato production. In stark contrast, the visible and near-infrared (Vis-NIR) spectroscopy technique emerges as an expedient solution (a non-destructive, expeditious, real-time, and cost-effective approach) to effectuate internal quality appraisal within the domain of agricultural product evaluation. Torres et al. (2015) used NIR reflectance spectroscopy to determine the SSC of Raf tomato based on modified PLSR (coefficient of determination for cross-validation is 0.75; standard error of prediction is 0.65%). Acharya et al. (2017) conducted a practical evaluation in the context of the non-destructive determination of the dry matter content of intact tomatoes (an indicator of the final mature SSC) using a handheld visible-short-wave NIR spectrophotometer. By using populations with different harvest dates or growth conditions for calibration and prediction, the dry matter prediction coefficient of determination (R^2) is 0.86-0.92, and the deviation is 0.14-0.03%. At different maturity stages of specific tomato varieties, Zhang et al. (2021) reported the acceptable prediction results of SSC evaluation by the self-developed Vis-NIR portable system (Rp = 0.70, RMSEP = 0.26%) and NIR integrating sphere system (Rp = 0.82, RMSEP = 0.21%). Aiming at the characteristics of tomato internal heterogeneous structure, in order to obtain more internal information of tomato as much as possible, Wang et al. (2018) built a tomato Vis-NIR diffuse transmission detection system to detect the SSC of cherry tomato, showing good prediction results (Rp = 0.93, RMSEP = 0.36%). However, the typical Vis-NIR spectroscopy is limited to a small area of measurement, and the spatial information of the sample obtained by single point measurement is limited. Liu et al. (2019) developed a dynamic online sorting system based on Vis-NIR diffuse transmission, and the sorting accuracy of SSC reached 91%. Yang Y. et al. (2022) based on the Vis-NIR diffuse transmission system, optimized the detection settings such as light path and light intensity, and compensated the model according to the height and weight physiological traits of tomato samples, and achieved good results (Rp = 0.91, RMSEP = 0.17%). The limitations of traditional single-point Vis-NIR measurement technology can be overcome by using on-line full transmission measurement and continuous data acquisition. Compared with the diffuse transmission mode, the full transmission mode and tomato multi-point spectral measurement acquisition can achieve a comprehensive characterization of the entire tomato information.

The main purpose of this study is to determine the best model for SSC prediction of tomato based on full transmission Vis-NIR spectroscopy detection technology. The specific purposes are as follows: (1) Collecting Vis-NIR spectral data of all tomato samples using full transmission Vis-NIR online detection equipment; (2) Processing continuous multi-point spectral data using the weighted average method; (3) Establishing a PLSR model based on full-

spectrum data, comparing the model's performance, and selecting the optimal preprocessing method and the best detection orientation; (4) Applying the least angle regression method to extract characteristic wavelengths in tomato SSC detection, and determining the optimal prediction model by combining prediction accuracy and stability.

2 Materials and methods

2.1 Experimental samples

In this study, we focused on two prominent tomato varieties, namely 'Provence' and 'Jingcai No.8' tomatoes, both of which enjoy substantial popularity in China. The 'Provence' tomato exhibits a thin skin, with succulent flesh, and attains a rich ruddy hue when reaching maturity. On the other hand, 'Jingcai No.8' tomato often referred to as strawberry tomato, features an orange-red or red peel with green shoulder, and its skin possesses a slight thickness. A comprehensive set of tomato samples encompassing various stages of maturity was meticulously collected to bolster the robustness of our predictive model for tomato SSC. These tomatoes were harvested from a farm located in the Fangshan District of Beijing, China. Tomato samples were collected from three maturity stages of half-ripe, hard-ripe and full-ripe, with a ratio of 1: 1: 1. The representative tomato samples obtained during this collection process are visually depicted in Figure 1. Moreover, in order to mitigate potential temperature-induced fluctuations that could influence the precision of our prediction model, the harvested tomatoes were transported to our laboratory facility and placed for a 24-hour period at a temperature of 20°C, with a relative humidity level of 60%, prior to the acquisition of spectral and SSC data (Yang X. et al., 2022).

To ensure the rigor and objectivity of our predictive model, we employed a systematic approach to partition all collected samples into two distinct subsets: a calibration set and a prediction set. This division was executed following a consistent ratio of 7:3, wherein 70% of the samples were allocated to the calibration set, responsible for the construction of the prediction model, while the remaining 30% constituted the prediction set, serving as an independent dataset for the assessment of model performance. And in order to mitigate the influence of random variability in sample partitioning and to provide a robust evaluation of our predictive model's efficacy, we undertook a systematic randomization process. Specifically, we repeated the sample division procedure 100 times, each time generating a new partition of samples. Subsequently, we constructed a predictive model based on the results of each of these 100 divisions. The culmination of these 100 modeling outcomes was then leveraged to derive an average, which serves as the foundational basis for evaluating the overall performance of our predictive model (Tian et al., 2022). This approach ensures a comprehensive and reliable assessment of the model's predictive capabilities.

2.2 Full transmission spectrum and real SSC acquisition

The full transmission Vis-NIR detection system, developed by the Intelligent Equipment Research Center of Beijing Academy of Agriculture and Forestry Sciences (Beijing, China), was used to acquire spectral data for all samples. This system, as depicted in Figure 2A, primarily comprises a highly sensitive spectrometer covering a wavelength range from 560 to 1072 nm and offering a spectral resolution of 0.5 nm. Furthermore, it is equipped with a conveyor platform featuring adjustable speed, a position sensor, and an illumination device consisting of a reflective halogen lamp (FUJI, JCR, 150W, 15V, Tokyo, Japan) with a focusing lens. The system is fortified with a shielding mechanism to prevent stray light

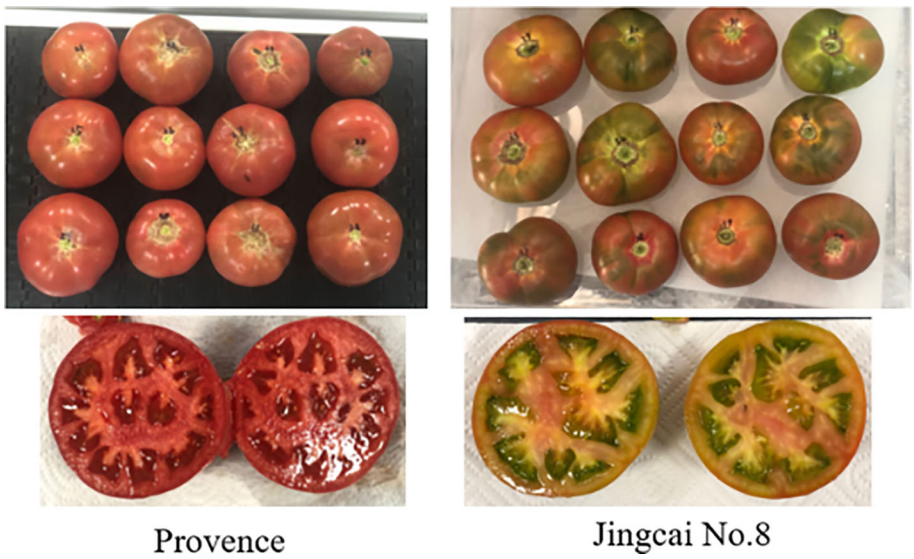


FIGURE 1
Tomato samples and cross sections.

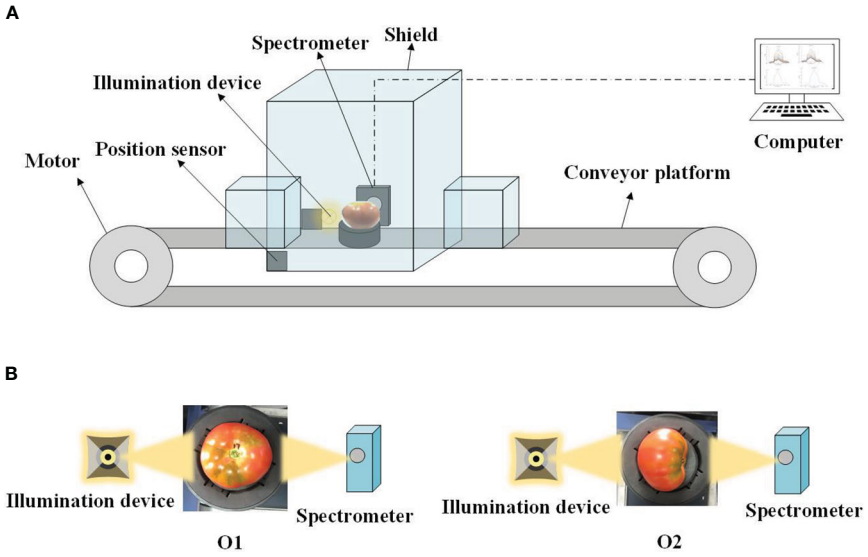


FIGURE 2
(A) Online full transmission spectroscopy and **(B)** detection orientations.

interference and is under the control of a computer-based system. Both the illumination device and the spectrometer are positioned on opposite sides of the conveyor belt.

To assess the influence of spectral measurement orientation on the accuracy of online SSC detection in tomatoes, spectral data for all tomato samples were collected in two orientations known as O1 and O2. In the O1 orientation, the tested tomato's stem-calyx axis was oriented perpendicular to the conveyor belt, with the stem facing upward. The sample received illumination from a halogen lamp at the equatorial position and was detected by the spectrometer on the opposite side. Conversely, in the O2 orientation, the stem-calyx axis of the tested tomato was parallel to the conveyor belt, with the stem directed towards the spectrometer. Schematic representations of these two detection orientations can be seen in Figure 2B. During each spectral measurement, the tomato sample was initially positioned on a fruit tray and moved at a consistent speed of 0.8 m/s. After the tomato sample passes through the sensor and the specified delay time, the spectrometer (integration time is 5ms) begins to continuously collect the spectral signals at each position on the sample. The multi-point spectra of each sample are roughly: 'Provence' tomato: 17-31 spectral curves; 'Jingcai No.8' tomato: 20-29 spectral curves.

Tomato SSC determination reference NY/T 2637-2014, using refractometer method, measuring instrument for digital Abbe refractometer. SSC measurements were performed immediately using the traditional method of destruction. Each complete tomato sample was first cut into pieces and squeezed into tomato juice in a wall-breaking machine. Then filter the tomato juice with gauze and squeeze into the beaker. After full shaking, the tomato juice was dropped on a hand-held digital refractometer, and the SSC value was manually recorded. Each measurement was repeated three times, and the average value was taken as the experimental value.

2.3 Data pre-processing

2.3.1 Multi-point spectrum weighted average

When the online full transmission mode collects the spectral signal of the sample, the incomplete signal is collected at both ends, so they are first eliminated before further data processing. The use of a weighting method signifies an effective strategy for spectral analysis (Somers et al., 2011; Zhu et al., 2019). In this study, distinct weights are assigned to individual data acquisition locations based on either the signal-to-noise ratio (*SNR*) (quality assessment metric). The spectrometer is placed in an environment without the sample to be tested, and the spectral signal in the environment is recorded to obtain the background. The operational process can be outlined as follows:

Firstly, an evaluation of the *SNR* is performed, enabling the quantitative characterization of *SNR* for each data acquisition point. *SNR* serves as a quantitative metric for signal quality assessment. Following this, weight factors are calculated, with each data acquisition point being assigned a weight based on the *SNR*, also referred to as the quality assessment metric. This study employs an inverse relationship where the weight factor increases with a higher quality assessment metric. In conclusion, a weighted average is carried out, involving the multiplication of spectral data associated with each acquisition point by its respective weight factor. This shows the derivation process of the final effective spectral representation. Through the application of the weighting method, high-quality spectral data significantly impacts the final effective spectrum, while the influence of low-quality spectral data is minimized.

2.3.2 Pretreatment of spectral data

To improve the correlation between tomato transmittance spectra and SSC and reduce the impact of unwanted signals and noise, it is customary to conduct preprocessing on the raw spectra.

Among the common preprocessing methods, the application of Savitzky-Golay smoothing (SG) is instrumental in making spectral curves more conducive to the recognition of features and localization of peaks, thus enhancing precision (Zhao et al., 2022). Furthermore, the use of standard normal variable transformation (SNV) serves to emphasize the positions of spectral peaks, streamlining the analysis of both spectral shape and peak locations (Bázár et al., 2016). In this study's context, two specific preprocessing techniques, a 13-point SG and SNV, have been implemented to refine the spectral data.

2.4 Prediction model and evaluation

An effective method for multivariate data analysis frequently employed in spectral analysis is PLSR. In this research, we constructed a PLSR model to delineate the quantitative relationship between the spectral matrix (**X**) and the matrix of SSC values in tomatoes (**Y**). To evaluate the mathematical approach's performance, we utilized metrics such as the calibration correlation coefficient (Rc), root mean square error of calibration (RMSEC), prediction correlation coefficient (Rp), and root mean square error of prediction (RMSEP). The specific calculation formula can be seen in Equations 1, 2. A robust model demonstrates correlation coefficients approaching 1 and lower root mean square error values (Li L. et al., 2022; Tian et al., 2023). In the process of model development, the optimal number of latent variables (LVs) is a critical consideration to prevent underfitting or overfitting issues (Diniz et al., 2015). To address this concern, we adopted a 5-fold cross-validation approach to determine the ideal number of LVs based on the minimum root mean square error of cross-validation (Li et al., 2023). The model was constructed using Matlab 2022b (Mathworks, Natick, MA).

$$\mathrm{RC}\ ,\mathrm{RP} = \sqrt{1 - \frac{\sum_{i=1}^{n}(y_i - \hat{y}_i)^2}{\sum_{i=1}^{n}(y_i - \overline{y}_i)^2}} \tag{1}$$

$$\mathrm{RMSEC}\ ,\mathrm{RMSEP} = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(y_i - \hat{y}_i)^2} \tag{2}$$

In this context, y_i and $\hat{y}_i$ denote the measured value and the predicted value for the ith tomato sample within either the calibration or prediction dataset. Additionally, $\overline{y}$ represents the mean value of the measured values for samples within the calibration or prediction dataset, and n signifies the total number of samples in either the calibration or prediction dataset.

2.5 Wavelength selection methods

Within the domain of full spectrum modeling, a total of 2047 wavelengths are present, including a substantial number of irrelevant and collinear variables. These extraneous wavelengths, in addition to complicating the model, have the potential to introduce interference, which could result in a reduction in model accuracy (Luo et al., 2022). Thus, to address this, the least angle regression (LARS) technique was implemented for the purpose of identifying and selecting pertinent wavelengths.

Efron et al. (2004) introduced the LARS algorithm, a method that functions as a feature selection technique applicable to both linear regression and sparse regression. Its primary aim is to pinpoint features with strong correlations to the response variables (SSC) and to retain only these essential features within the model. This approach effectively simplifies the model, thus enhancing its capacity for generalization. LARS proceeds by incrementally integrating features and moving along the gradient direction of these features in each step. What sets LARS apart is its utilization of the regression coefficient path, allowing the simultaneous addition of multiple features. At the core of LARS is the consistent alignment with the prevailing gradient direction at each step, coupled with the allocation of suitable step sizes between features. This approach enables LARS to promptly and reliably identify characteristic wavelengths highly correlated with SSC. Nonetheless, LARS may face efficiency challenges when handling extensive feature sets. To tackle this limitation, the present study introduces regularization terms in the form of the L1 norm (Lasso penalty) and L2 norm (Ridge penalty). The L1 norm penalty streamlines feature selection by reducing coefficients of irrelevant features to zero, resulting in the construction of a sparse model. This feature is particularly advantageous when dealing with high-dimensional data and problems involving the selection of essential wavelengths for SSC analysis. Conversely, the L2 norm penalty trims model parameters to prevent overfitting and enhance model generalization. Differing from L1 regularization, L2 regularization refrains from entirely zeroing out parameters, offering controlled adjustment of model complexity. This feature proves beneficial in addressing issues related to collinearity and augmenting model stability.

In this study, the specific regularization term is denoted by λ. A range of λ values is systematically selected, typically starting with a smaller value and gradually increasing it. Model performance is monitored, and an appropriate λ value is selected accordingly. Characteristic wavelength selection was carried out using Matlab 2022b (Mathworks, Natick, MA).

3 Results and discussion

3.1 SSC values of all samples

In this experiment, tomato samples with different maturity were collected to establish SSC prediction model. The SSC values of all samples measured are shown in Table 1. The SSC range of 'Provence' tomato is 3.8-8.7°Brix, and that of 'Jingcai No.8' tomato is 4.5-9.8° Brix. The standard deviations (SD) were 1.1 and 1.2°Brix, respectively. SSC has a wide range of distribution. Combined with the characteristics of different maturity samples, a more comprehensive and accurate SSC prediction model can be established. This method can enhance the robustness of the model and improve the accuracy of the prediction results (Li Y. et al., 2022).

TABLE 1 SSC values (°Brix) of tomato samples.

Variety	Range	No. of samples	Mean	SD
Provence	3.8-8.7	92	5.8	1.1
Jingcai NO.8	4.5-9.8	96	7.4	1.2

The SSC values (°Brix) were measured by refractometer.

At the same time, it can also better cope with the different maturity of tomato samples that may occur in practical applications.

3.2 Analysis of tomato spectral feature

Figure 3A shows that the multi-point spectral curves collected by each sample have problems such as low *SNR* and intensity supersaturation due to the acquisition method of online full transmission measurement, which may be caused by the texture color characteristics of tomatoes and the online acquisition method of spectrum. According to the spectral contribution of different parts, we can see the curve shown in Figure 3B. After the weighted average method, the noise in the data is effectively reduced, the *SNR* is improved, and the characteristic peak is enhanced, which is more conducive to the establishment of the subsequent prediction model. In Figure 3C, we can observe that the spectral curve characteristics of the same variety in different directions are basically similar. The main difference is the intensity, which may be due to the influence of the internal cavity structure of the sample on the propagation light path. Because the optical path of O2 orientation is simpler, the optical path distance is shorter, and the flesh tissue is less penetrated, the spectral curves of both varieties show that the intensity of O2 orientation is higher than that of O1.

3.3 Prediction of tomato SSC with full spectra

Table 2 presents the outcomes derived from PLSR modeling applied to spectral data with the utilization of diverse preprocessing techniques. Notably, whether considering the 'Provence' tomato or 'Jingcai No.8' tomato, the results consistently manifest superior O2 performance as opposed to O1. This intriguing phenomenon is likely attributed to the inherent simplicity of the optical propagation pathway or the shortened propagation distance in the O2 direction. Conversely, the trajectory of incident light in the O1 direction necessitates traversal through discrete cavities, often yielding a more intricate optical pathway and extended propagation distances. In this study, the application of the SG method yields results marginally less favorable than the unprocessed data. This discrepancy could be ascribed to the potential obfuscation of essential spectral features by the SG method. Enhanced results are attainable through the adoption of preprocessing methodologies such as SNV. SNV preprocessing methods adeptly ameliorate scattering influences within the spectra, thereby endowing the data with increased stability, consistency, and diminished variability, consequently yielding a positive influence on modeling and quantitative analytical outcomes. The findings indicate that, subsequent to optimal preprocessing, the samples exhibit Rp = 0.81, RMSEP = 0.69°Brix ('Provence' tomato) and Rp = 0.84, RMSEP = 0.64°Brix ('Jingcai No.8' tomato). Obviously, the SNV preprocessing process magnifies the spectral attributes, making the spectrum clearer, more consistent and more prominent, thereby improving the quality of the spectral data set. In summary, this study established a robust PLSR model for SSC prediction. The model is founded on the determination of the optimal detection orientation (O2) in synergy with the implementation of the most effective preprocessing method (SNV).

3.4 Determination of the optimal model

While the full-spectrum PLSR model can effectively predict SSC quantitatively, most full-spectrum models exhibit instability due to notable disparities between Rc and Rp. Given that an excessive number of spectral variables employed in modeling can lead to overfitting, this study implemented characteristic wavelength selection to optimize the model. Through the selection of wavelengths, superfluous features are reduced, rendering the model more concise and efficient. Since O2 represents the optimal detection orientation, and SNV serves as the most effective preprocessing method, variable selection was exclusively based on the spectral data acquired in the O2 orientation and after preprocessing using SNV.

During the deployment of the LARS method for characteristic wavelength selection, initialization is initiated at the outset. The model begins by not selecting any characteristic wavelengths, with all coefficients set to zero. Subsequently, in each step, the system identifies the characteristic wavelength displaying the highest correlation with SSC and calculates the absolute value of this correlation. The selection of characteristic wavelengths follows a path along the minimum angle. Then, the coefficients are updated, with the coefficient of the selected characteristic wavelength gradually increasing until its correlation with another characteristic wavelength equals it. The regularization parameter λ is then progressively adjusted to achieve a balance between characteristic wavelength selection and model sparsity. Ultimately, the steps following the initializations are reiterated until the selection outcome converges.

A series of selected λ values and the corresponding Rp relationship are shown in Figure 4. It can be seen that LARS-L1 and LARS-L2 are significantly different in the selection range of λ,

10.3389/fpls.2024.1324753

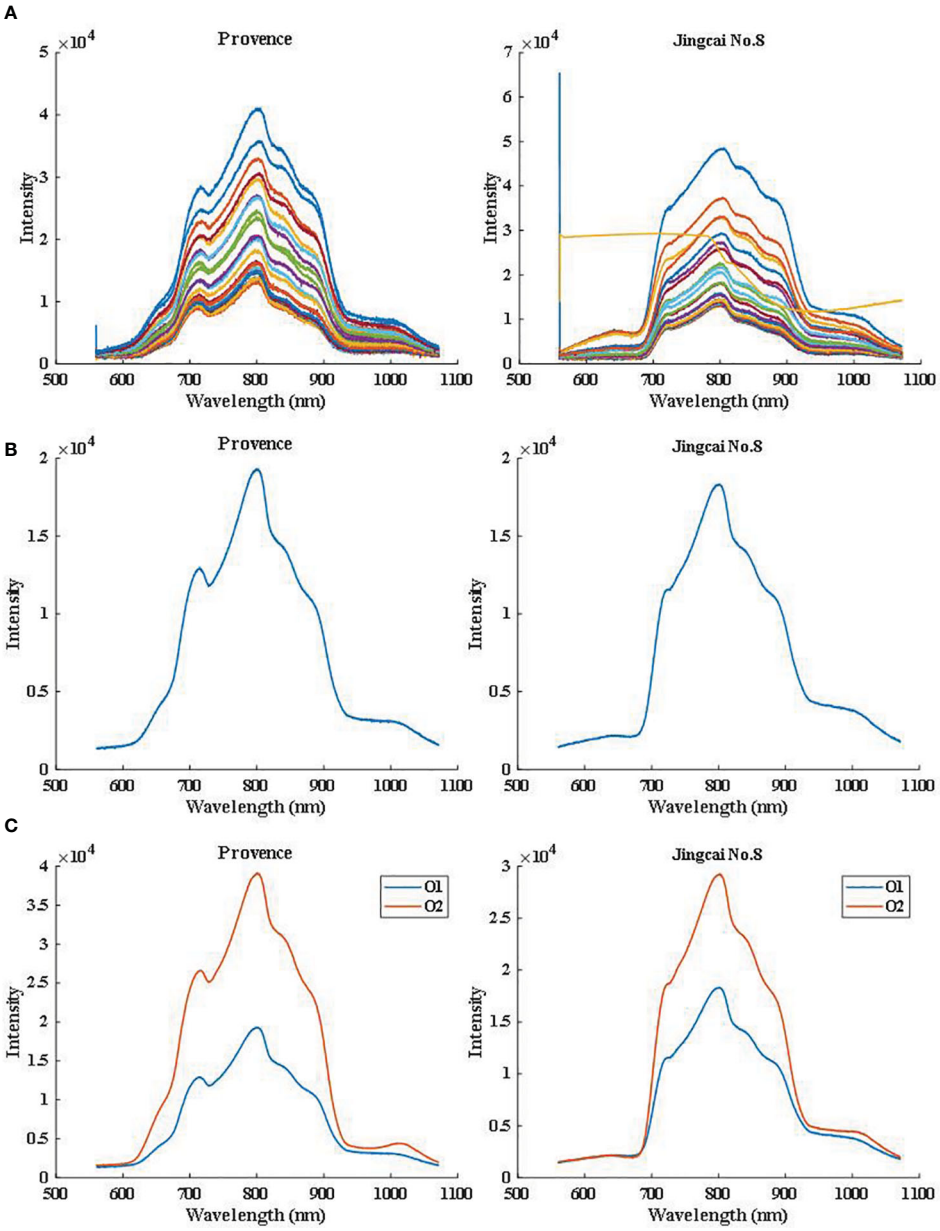


FIGURE 3
Spectral curve from **(A)** multi-point raw data, **(B)** weighted average and **(C)** two detection orientations.

TABLE 2 Prediction results of SSC of two varieties of tomatoes based on PLSR model established using full-spectrum data combined with different detection orientations and preprocessing methods.

Varieties	Orientations	Methods	LVs	Rc	RMSEC	Rp	RMSEP
Provence	O1	RAW	11	0.91	0.44	0.66	0.89
		SG	11	0.88	0.51	0.65	0.88
		SNV	9	0.95	0.31	0.70	0.84
	O2	RAW	10	0.89	0.50	0.73	0.82
		SG	11	0.88	0.52	0.74	0.80
		SNV	**9**	**0.94**	**0.36**	**0.81**	**0.69**

(Continued)

TABLE 2 Continued

Varieties	Orientations	Methods	LVs	Rc	RMSEC	Rp	RMSEP
Jingcai NO.8	O1	RAW	9	0.94	0.39	0.62	1.04
		SG	11	0.89	0.53	0.62	1.10
		SNV	10	0.98	0.25	0.79	0.73
	O2	RAW	10	0.91	0.47	0.74	0.81
		SG	10	0.88	0.55	0.74	0.81
		SNV	**9**	**0.96**	**0.31**	**0.84**	**0.64**

which may be due to the encouragement of L1 regularization to sparsity: L1 regularization encourages sparsity by punishing the absolute value of the coefficient, that is, encouraging the model to reduce the coefficient of most features to zero, so as to select the most important features, which usually requires a relatively small λ to achieve. L2 regularization encourages smoothness: L2 regularization reduces the magnitude of the coefficients by penalizing the square of the coefficients, thereby encouraging the coefficients of the feature to be evenly distributed, but it does not compress the coefficients to zero. Therefore, in order to achieve an effective L2 regularization effect, a relatively large λ is usually required. Although there are some fluctuations in the model Rp results obtained under different λ values, the changes are also around 0.01-0.03, and the results are relatively good (Rp > 0.90).

In the prediction model results of LARS-L1 method, the Rp of 'Provence' tomato and 'Jingcai No.8' tomato were both above 0.93. When the regularization parameter λ = 0.004, the effect was the best ('Provence' tomato: Rc = 0.98, RMSEC = 0.23°Brix, Rp = 0.95, RMSEP = 0.35°Brix; 'Jingcai No.8' tomato: Rc = 0.98, RMSEC = 0.20°Brix, Rp = 0.96, RMSEP = 0.33°Brix). L1 regularization tends

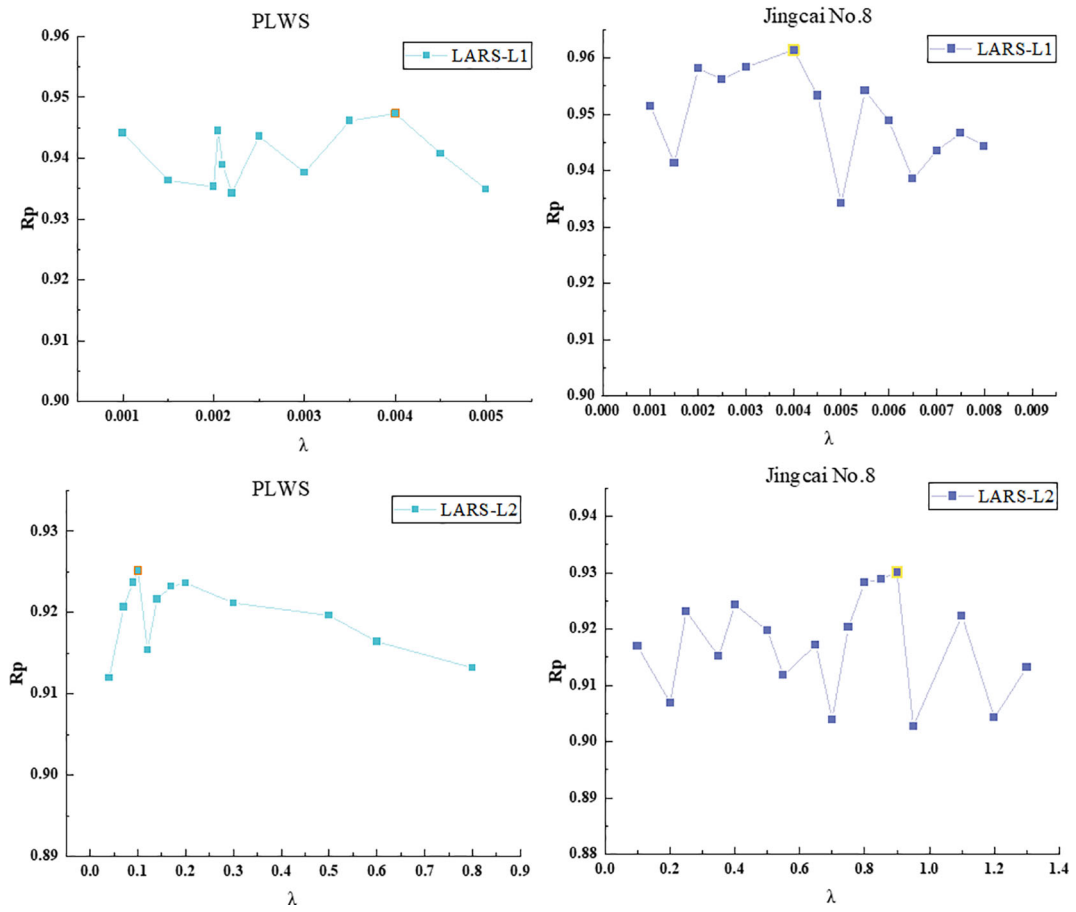


FIGURE 4
The process of Rp changing with λ in the process of characteristic wavelength selection.

10.3389/fpls.2024.1324753

to generate sparse solutions, so that some weights are zero, and a small number of characteristic wavelengths that have a significant impact on SSC prediction can be selected to filter out wavelengths that are not important for prediction. The best performance of the LARS-L1 method is due to its sparseness of L1 regularization and better capture of the correlation between wavelength and SSC. LARS-L2 obtained the best effect of 'Provence' tomato at $\lambda = 0.1$ (Rc = 0.97, RMSEC = 0.26°Brix, Rp = 0.93, RMSEP = 0.42°Brix). 'Jingcai No.8' tomato achieved the best results at $\lambda = 0.9$ (Rc = 0.97, RMSEC = 0.28°Brix, Rp = 0.93, RMSEP = 0.45°Brix). The penalty of L2 regularization on feature weights is balanced, and the weight of ownership is relatively evenly reduced without deleting some features too much. When there are some relatively weak wavelengths in the data that still contribute to the prediction, L2 regularization preserves these wavelengths. This may be the reason that the performance of LARS-L2 is slightly lower than that of LARS-L1. The LARS-L1 and LARS-L2 methods have less influence on the correlation of features because they constrain the feature weights through regularization and exhibit a degree of stability. The best results of the two characteristic wavelength selection methods for the two varieties were placed in Table 3.

In practical applications, aside from predictive accuracy, model stability is also a crucial consideration. Figure 5 illustrates the error bar chart for the prediction results of SSC for two tomato varieties based on 100 modeling iterations, incorporating the optimal feature wavelength selection from two methods. Each data point in the figure is associated with an error bar, and the central mark denotes the mean value, reflecting the data's central tendency. These error bars signify the data's dispersion or uncertainty. It is evident that the results of both methods align with the trends depicted in Figure 4. The LARS-L1 method exhibits the highest mean Rp value, consistent with the central tendency of the data, and features shorter error bars, indicative of lower dispersion. Therefore, overall, the model demonstrates a heightened level of stability.

TABLE 3 SSC prediction results obtained by characteristic wavelength PLSR models.

Varieties	Wavelength selection methods	λ	LVs	No. of variables	Rc	RMSEC	Rp	RMSEP
Provence	**LARS-L1**	**0.004**	**9**	**29**	**0.98**	**0.23**	**0.95**	**0.35**
	LARS-L2	0.1	10	53	0.97	0.26	0.93	0.42
Jingcai NO.8	**LARS-L1**	**0.004**	**9**	**63**	**0.98**	**0.20**	**0.96**	**0.33**
	LARS-L2	0.9	11	45	0.97	0.28	0.93	0.45

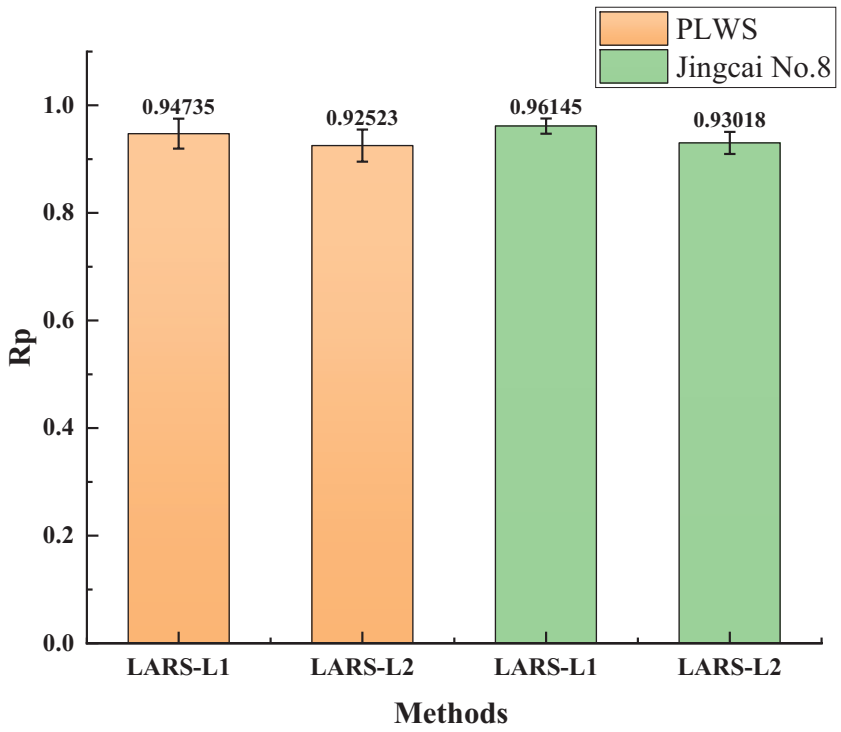


FIGURE 5
The error bar graph of 100 iterations under optimal λ.

4 Conclusions

In this study, the SSC of tomato was successfully determined using the full transmission Vis-NIR spectroscopy online detection equipment. The results show that the weighted average method can improve the spectral quality of multi-point spectral data. The prediction performance of O2 is better than that of O1 in the detection orientation, and the prediction performance of the full-spectrum PLSR model constructed after SNV pretreatment is significantly improved. For the samples in the prediction set, the results of the two varieties of tomatoes were 'Provence' tomato: Rp = 0.81, RMSEP = 0.69° Brix; 'Jingcai No.8' tomato: Rp = 0.84, RMSEP = 0.64°Brix. In addition, in order to reduce the number of variables involved in modeling, the LARS method combined with L1 and L2 regularization is used to select the characteristic wavelengths to construct the PLSR model. The results show that the prediction accuracy of the characteristic wavelength selection model is better than that of the full spectrum model. When λ was set to 0.004, the characteristic wavelengths selected by the LARS-L1 method achieved the best results on the SSC prediction models of the two varieties of tomatoes ('Provence' tomato: Rp = 0.95, RMSEP = 0.35° Brix; 'Jingcai No.8' tomato: Rp = 0.96, RMSEP = 0.33°Brix). Under the condition of optimal λ, 100 modeling calculations were carried out to further verify the stability of the model. Finally, O2-SNV-LARS-L1-PLSR was determined as the best model for quantitative detection of tomato SSC, and it showed that this method combined with full transmission Vis-NIR spectroscopy had the potential for non-destructive detection of SSC in tomato.

Data availability statement

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

Ethics statement

This article has no any study with human participants or animals by any of the authors.

Author contributions

SL: Writing – original draft. JL: Writing – original draft. QW: Writing – review & editing. RS: Writing – review & editing. XY: Funding acquisition, Writing – review & editing. QZ: Supervision, Writing – review & editing.

Funding

The author(s) declare financial support was received for the research, authorship, and/or publication of this article. This work was supported by the Science and Technology Innovation Ability Construction Project of Beijing Academy of Agriculture and Forestry Sciences (Project No. KJCX20240503), the National Natural Science Foundation of China (32360605), Bingtuan scientific and technological innovation talent plan (2021CB042, 2023CB016).

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 author(s) declared that they were an editorial board member of Frontiers, at the time of submission. This had no impact on the peer review process and the final decision.

Publisher's note

References

Acharya, U. K., Subedi, P. P., and Walsh, K. B. (2017). Robustness of tomato quality evaluation using a portable Vis-SWNIRS for dry matter and colour. *Int. J. Anal. Chem* 2017, 2863454. doi: 10.1155/2017/2863454

Ali, M. Y., Sina, A. A. I., Khandker, S. S., Neesa, L., Tanvir, E. M., Kabir, A., et al. (2020). Nutritional composition and bioactive compounds in tomatoes and their impact on human health and disease: A review. *Foods* 10 (1), 45. doi: 10.3390/foods10010045

Bázár, G., Romvári, R., Szabó, A., Somogyi, T., Éles, V., and Tsenkova, R. (2016). NIR detection of honey adulteration reveals differences in water spectral pattern. *Food Chem.* 194, 873–880. doi: 10.1016/j.foodchem.2015.08.092

Diniz, P. H. G. D., Pistonesi, M. F., and Araújo, M. C. U. (2015). Using i SPA-PLS and NIR spectroscopy for the determination of total polyphenols and moisture in commercial tea samples. *Anal. Methods* 7 (8), 3379–3384. doi: 10.1039/C4AY03099K

Efron, B., Hastie, T., Johnstone, I., and Tibshirani, R. (2004). Least angle regression. *Ann. Stat.* 32 (2), 407–451.

Gómez, A. H., Wang, J., Hu, G., and Pereira, A. G. (2008). Monitoring storage shelf life of tomato using electronic nose technique. *J. Food Eng.* 85 (4), 625–631. doi: 10.1016/j.jfoodeng.2007.06.039

Guan, Z., Biswas, T., and Wu, F. (2018). *The US tomato industry: An overview of production and trade: FE1027, 9/2017* Vol. 2018 (EDIS).

Li, J., Lu, Y., and Lu, R. (2023). Detection of early decay in navel oranges by structured-illumination reflectance imaging combined with image enhancement and segmentation. *Postharvest Biol. Technol.* 196, 112162.

Li, L., Huang, W., Wang, Z., Liu, S., He, X., and Fan, S. (2022). Calibration transfer between developed portable Vis/NIR devices for detection of soluble solids contents in apple. *Postharvest Biol. Technol.* 183, 111720. doi: 10.1016/j.postharvbio.2021.111720

Li, Y., Ma, B., Li, C., and Yu, G. (2022). Accurate prediction of soluble solid content in dried Hami jujube using SWIR hyperspectral imaging with comparative analysis of models. *Comput. Electron. Agric.* 193, 106655. doi: 10.1016/j.compag.2021.106655

Liu, Y. D., Rao, Y., Sun, X. D., Xiao, H. C., Jiang, X. G., Zhu, K., et al. (2019). The online detection model research of tomatoes' bruise and SSD. *Spectr. Spectral Anal.* 39 (12), 3910–3915. doi: 10.3964/j.issn.1000-0593(2019)12-3910-06

Luo, W., Tian, P., Fan, G., Dong, W., Zhang, H., and Liu, X. (2022). Non-destructive determination of four tea polyphenols in fresh tea using visible and near-infrared spectroscopy. *Infrared Phys. Technol.* 123, 104037. doi: 10.1016/j.infrared.2022.104037

Mei, M., and Li, J. (2023). An overview on optical non-destructive detection of bruises in fruit: Technology, method, application, challenge and trend. *Comput. Electr. Agricul.* 213, 108915.

Mollazade, K., Omid, M., Akhlaghian Tab, F., Rezaei Kalaj, Y., and Mohtasebi, S. S. (2015). Data mining-based wavelength selection for monitoring quality of tomato fruit by backscattering and multispectral imaging. *Int. J. Food Properties* 18 (4), 880–896. doi: 10.1080/10942912.2013.835822

Najjar, K., and Abu-Khalaf, N. (2021). Non-destructive quality measurement for three varieties of tomato using VIS/NIR spectroscopy. *Sustainability* 13 (19), 10747. doi: 10.3390/su131910747

Nikbakht, A. M., Tavakkoli, H. T., Malekfar, R., and Gobadian, B. (2011). Nondestructive determination of tomato fruit quality parameters using raman spectroscopy. *J. Agricultural Sci. Tech.* 13 (4), 517–526.

Passam, H. C., Karapanos, I. C., Bebeli, P. J., and Savvas, D. (2007). A review of recent research on tomato nutrition, breeding and post-harvest technology with reference to fruit quality. *Eur. J. Plant Sci. Biotechnol.* 1 (1), 1–21.

Perveen, R., Suleria, H. A. R., Anjum, F. M., Butt, M. S., Pasha, I., and Ahmad, S. (2015). Tomato (Solanum lycopersicum) carotenoids and lycopenes chemistry; metabolism, absorption, nutrition, and allied health claims—A comprehensive review. *Crit. Rev. Food Sci. Nutr.* 55 (7), 919–929. doi: 10.1080/10408398.2012.657809

Ponce-Valadez, M., Escalona-Buendía, H. B., Villa-Hernández, J. M., de León-Sánchez, F. D., Rivera-Cabrera, F., Alia-Tejacal, I., et al. (2016). Effect of refrigerated storage (12.5 C) on tomato (Solanum lycopersicum) fruit flavor: A biochemical and sensory analysis. *Postharvest Biol. Technol.* 111, 6–14. doi: 10.1016/j.postharvbio.2015.07.010

Rahman, A., Kandpal, L. M., Lohumi, S., Kim, M. S., Lee, H., Mo, C., et al. (2017). Nondestructive estimation of moisture content, pH and soluble solid contents in intact tomatoes using hyperspectral imaging. *Appl. Sci.* 7 (1), 109. doi: 10.3390/app7010109

Salehi, B., Sharifi-Rad, R., Sharopov, F., Namiesnik, J., Roointan, A., Kamle, M., et al. (2019). Beneficial effects and potential risks of tomato consumption for human health: An overview. *Nutrition* 62, 201–208. doi: 10.1016/j.nut.2019.01.012

Ścibisz, I., Reich, M., Bureau, S., Gouble, B., Causse, M., Bertrand, D., et al. (2011). Mid-infrared spectroscopy as a tool for rapid determination of internal quality parameters in tomato. *Food Chem.* 125 (4), 1390–1397. doi: 10.1016/j.foodchem.2010.10.012

Skolik, P., McAinsh, M. R., and Martin, F. L. (2019). ATR-FTIR spectroscopy non-destructively detects damage-induced sour rot infection in whole tomato fruit. *Planta* 249, 925–939. doi: 10.1007/s00425-018-3060-1

Somers, B., Asner, G. P., Tits, L., and Coppin, P. (2011). Endmember variability in spectral mixture analysis: A review. *Remote Sens. Environ.* 115 (7), 1603–1616. doi: 10.1016/j.rse.2011.03.003

Tian, X., Cheng, L. P., Wang, Q. Y., Li, J. B., Yang, Y., Fan, S. X., et al. (2022). Optimization of online determination model for sugar in a whole apple using full transmittance spectrum. *Spectrosc. Spectral Anal.* 42 (6), 1907–1914. doi: 10.3964/j.issn.1000-0593(2022)06-1907-08

Tian, X., Liu, X., He, X., Zhang, C., Li, J., and Huang, W. (2023). Detection of early bruises on apples using hyperspectral reflectance imaging coupled with optimal wavelengths selection and improved watershed segmentation algorithm. *J. Sci. Food Agricul.* 103 (13), 6689–6705.

Torres, I., Pérez-Marín, D., de la Haba, M. J., and Sánchez, M. T. (2015). Fast and accurate quality assessment of Raf tomatoes using NIRS technology. *Postharvest Biol. Technol.* 107, 9–15. doi: 10.1016/j.postharvbio.2015.04.004

Wang, F., Peng, Y. K., Tang, X. Y., Li, L., and Li, Y. (2018). Near infrared nondestructive testing of soluble solids content of cherry tomato. *J. Chin. Institute Food Sci. Technol.* 18 (10), 235–240. doi: 10.11895/j.issn.0253-3820.181164

Yang, Y., Zhao, C., Huang, W., Tian, X., Fan, S., Wang, Q., et al. (2022). Optimization and compensation of models on tomato soluble solids content assessment with online Vis/NIRS diffuse transmission system. *Infrared Phys. Technol.* 121, 104050. doi: 10.1016/j.infrared.2022.104050

Yang, X., Zhu, L., Huang, X., Zhang, Q., Li, S., Chen, Q., et al. (2022). Determination of the soluble solids content in korla fragrant pears based on visible and near-infrared spectroscopy combined with model analysis and variable selection. *Front. Plant Sci.* 13, 938162. doi: 10.3389/fpls.2022.938162

Youssef, M. A., and Eissa, M. A. (2017). Comparison between organic and inorganic nutrition for tomato. *J. Plant Nutr.* 40 (13), 1900–1907. doi: 10.1080/01904167.2016.1270309

Zhang, D., Yang, Y., Chen, G., Tian, X., Wang, Z., Fan, S., et al. (2021). Nondestructive evaluation of soluble solids content in tomato with different stage by using Vis/NIR technology and multivariate algorithms. *Spectrochimica Acta Part A: Mol. Biomol. Spectrosc.* 248, 119139. doi: 10.1016/j.saa.2020.119139

Zhao, R., An, L., Tang, W., Gao, D., Qiao, L., Li, M., et al. (2022). Deep learning assisted continuous wavelet transform-based spectrogram for the detection of chlorophyll content in potato leaves. *Comput. Electron. Agric.* 195, 106802. doi: 10.1016/j.compag.2022.106802

Zhu, S., Cui, X., Xu, W., Chen, S., and Qian, W. (2019). Weighted spectral reconstruction method for discrimination of bacterial species with low signal-to-noise ratio Raman measurements. *RSC Adv.* 9 (17), 9500–9508. doi: 10.1039/C9RA00327D

Zhu, Q., He, C., Lu, R., Mendoza, F., and Cen, H. (2015). Ripeness evaluation of 'Sun Bright'tomato using optical absorption and scattering properties. *Postharvest Biol. Technol.* 103, 27–34. doi: 10.1016/j.postharvbio.2015.02.007

frontiers | Frontiers in Plant Science

TYPE Original Research
PUBLISHED 31 January 2024
DOI 10.3389/fpls.2024.1292365

OPEN ACCESS

EDITED BY
Naveen Kumar Mahanti,
Dr. Y.S.R. Horticultural University, India

REVIEWED BY
Shekh Mukhtar Mansuri,
Indian Council of Agricultural Research (ICAR), India
Konga Upendar,
Centurion University, India
Jiangbo Li,
Beijing Academy of Agriculture and Forestry Sciences, China

*CORRESPONDENCE
Yandan Lin
ydlin@fudan.edu.cn

RECEIVED 11 September 2023
ACCEPTED 15 January 2024
PUBLISHED 31 January 2024

CITATION
Xu W, Wei L, Cheng W, Yi X and Lin Y (2024) Non-destructive assessment of soluble solids content in kiwifruit using hyperspectral imaging coupled with feature engineering. *Front. Plant Sci.* 15:1292365. doi: 10.3389/fpls.2024.1292365

Non-destructive assessment of soluble solids content in kiwifruit using hyperspectral imaging coupled with feature engineering

Wei Xu[1,2], Liangzhuang Wei[3], Wei Cheng[1], Xiangwei Yi[3] and Yandan Lin[1,2,3]*

[1]Institute for Electric Light Sources, School of Information Science and Technology, Fudan University, Shanghai, China, [2]Institute for Six-sector Economy, Fudan University, Shanghai, China, [3]Academy for Engineering & Technology, Fudan University, Shanghai, China

The maturity of kiwifruit is widely gauged by its soluble solids content (SSC), with accurate assessment being essential to guarantee the fruit's quality. Hyperspectral imaging offers a non-destructive alternative to traditional destructive methods for SSC evaluation, though its efficacy is often hindered by the redundancy and external disturbances of spectral images. This study aims to enhance the accuracy of SSC predictions by employing feature engineering to meticulously select optimal spectral features and mitigate disturbance effects. We conducted a comprehensive investigation of four spectral pre-processing and nine spectral feature selection methods, as components of feature engineering, to determine their influence on the performance of a linear regression model based on ordinary least squares (OLS). Additionally, the stacking generalization technique was employed to amalgamate the strengths of the two most effective models derived from feature engineering. Our findings demonstrate a considerable improvement in SSC prediction accuracy post feature engineering. The most effective model, when considering both feature engineering and stacking generalization, achieved an $RMSE_p$ of 0.721, a $MAPE_p$ of 0.046, and an RPD_p of 1.394 in the prediction set. The study confirms that feature engineering, especially the careful selection of spectral features, and the stacking generalization technique are instrumental in bolstering SSC prediction in kiwifruit. This advancement enhances the application of hyperspectral imaging for quality assessment, offering benefits that extend across the agricultural industry.

KEYWORDS

kiwifruit, soluble solids content, feature engineering, stacking generalization, spectral imaging

1 Introduction

Kiwifruit (Actinidia deliciosa) is a popular fruit known for its unique flavor and nutritional benefits. As a typical climacteric fruit, it continues ripening even after being harvested. This post-harvest ripening process makes kiwifruit highly perishable and requires careful handling and storage to maintain its quality. The assessment of its quality and maturity commonly relies on the measurement of soluble solids content (SSC). On the one hand, SSC serves as an indicator of the sugar content in kiwifruit, for sugars constituting approximately 81% of the total SSC (Tian et al., 2022). On the other hand, SSC exhibits a consistent pattern of variation over time in storage. Throughout the storage period, as time goes by, the starch and pectin present in the kiwifruit undergo hydrolysis, leading to a gradual increase in SSC. Therefore, monitoring the SSC of kiwifruit is effective for evaluating its quality and maturity. However, the determination of SSC, being an internal attribute of fruit, often involves destructive techniques like refractometry, which requires the extraction of juice or pulp from fruit. These methods are time-consuming, labor-intensive and cause damage to the fruit, preventing the repeated utilization of samples. Consequently, there is an increasing demand for non-destructive and expeditious techniques that can precisely estimate the SSC of kiwifruit.

Hyperspectral imaging has emerged as a promising non-destructive method for assessing the quality of various agricultural products (Yao et al., 2013; Huang et al., 2018). This technique enables the measurement of spectral reflectance across a broad range of wavelengths, providing detailed insights into the chemical and physical properties of samples. In the case of kiwifruit, the visible near-infrared (Vis-NIR) spectral range contains valuable information related to the absorption of O–H, N–H, and C–H vibrations (Guo et al., 2017; Xu et al., 2023). These vibrational modes facilitate the identification and quantification of key chemical constituents associated with SSC, such as sugars and other organic compounds. Through the employment of regression models, relevant information can be extracted from spectral reflectance, leading to the establishment of a strong relationship between the observed spectral features and SSC measurements. Once the regression model is constructed, predicting SSC becomes a straightforward process, allowing for the non-destructive estimation of SSC values (Nicolaï et al., 2007).

Various well-designed regression models, such as partial least squares regression (PLSR) (Lee et al., 2022), support vector machine regression (SVR) (Ma et al., 2018), and artificial neural network (ANN) (Pullanagari and Li, 2021) have been developed to establish the relationship between observed spectral features and SSC measurements. However, the high-dimensional nature of spectral features can pose challenges to regression models. These features often contain redundant information and are influenced by various disturbances (e.g., sample differences, environmental noise, and baseline drift). Excessive redundant information for regression models not only results in prolonged hardware and software runtime but also compromises the regression performance, leading to unreliable estimations of SSC values (Xiaobo et al., 2010).

Unlike previous research that focuses on refining regression or machine learning models, our study intentionally emphasizes the importance of eliminating redundancies and disturbances in the initial phase of model development to enhance SSC prediction for kiwifruit—a crucial yet frequently underestimated step in existing studies.

The quality and suitability of input features significantly influence the performance of regression models. Carefully selected features provide more relevant information, resulting in simpler models and improved results. Conversely, the inclusion of irrelevant features can negatively impact the model's ability to generalize. In contrast to complex models, which may present challenges in interpretation and fine-tuning, simpler models with more effective features tend to yield more reliable results (Xiaobo et al., 2010). Hence, it is essential to pay meticulous attention to the pre-processing and selection of these features. These tasks, involving data converting and filtering before model building, are collectively referred to as feature engineering. In general, feature engineering involves spectral pre-processing and selection to effectively mitigate the impact of various disturbances, eliminate irrelevant features, and identify the most informative ones. Its ultimate goal is to generate enhanced features that are well-suited for integration into regression models. By prioritizing the use of more effective features and employing simple models, we can strike a balance between model complexity and performance, thus leading to more accurate and interpretable regression results.

In this study, we focus on investigating the effectiveness of feature engineering in enhancing the performance of SSC prediction in kiwifruit using hyperspectral imaging. To achieve this goal, we employed a linear regression model based on ordinary least squares (OLS) due to its simplicity and interpretability. Subsequently, we conducted a systematic evaluation and comparison of the variations in the regression performance under different combinations of four spectral pre-processing methods and nine spectral feature selection methods (details will be provided in section 2.3~2.5). Through this comprehensive analysis, our study not only demonstrates the positive impact of feature engineering but also identifies the optimal condition that yields the best regression performance. Additionally, we introduce the stacking generalization technique to integrate the strengths of two best-performing models which are achieved through above feature engineering, thus effectively addressing overfitting issues, and further improving the regression performance. This study highlights the potential of feature engineering and the stacking generalization technique in SSC prediction for kiwifruit, providing practical insights for quality assessment in the kiwifruit industry. The application of these techniques holds promise for more efficient and reliable SSC prediction, benefiting the kiwifruit industry and potentially extending to other agricultural produce quality assessment domains.

2 Materials and methods

2.1 Preparation of kiwifruit samples

In June 2023, a substantial number of kiwifruit samples were obtained from an agricultural plantation situated in Shaanxi Province, China. Following the removal of unqualified samples such as unripe, overripe, or mechanically damaged ones, a total of 116 kiwifruit samples with intact skin were selected for utilization in this experiment.

Prior to conducting the spectral acquisition step, a meticulous wiping procedure was carried out using soft tissue paper to eliminate any lint present on the surface of kiwifruit samples. This step was taken to mitigate the potential influence of lint on the spectral acquisition step.

Immediately following the spectral acquisition step, the sample preparation for the SSC measurement was conducted under the guidelines of the NY/T 2637-2014 standard. This sample preparation entails peeling the samples along their equators, removing the pulp, and extracting the juice through pressing. The kiwifruit juice will be introduced into the detection tank of one refractometer for subsequent SSC measurement.

2.2 Spectral acquisition and SSC measurement

A custom-built hyperspectral imaging system is specifically developed to capture spectral images of the kiwifruit samples, consisting of four main components: a spectral imaging camera (Specim FX10, Konica Minolta, Inc., Japan), a motorized positioning sample platform, two halogen area light sources, and a computer installed with suitable data acquisition software (see Figure 1). Among them, the Specim FX10 spectral imaging camera provides a spectral resolution of 400 ~ 1000 nm (due to the low signal-to-noise ratio in the lower wavelength regions, only data from wavelengths above 450 nm were exclusively utilized in this study) and works in a push-broom mode, thus necessitating a motorized positioning sample platform. To ensure an accurate aspect ratio in the captured spectral images, it is crucial to carefully adjust the advancing speed of the platform and the exposure time of the spectral imaging camera to match each other. The two light sources were positioned symmetrically to uniformly illuminate the camera's field of view. This arrangement aims to ensure consistent spectral response across different positions within the imaged area. For stable and accurate measurements, a one-hour warm-up and black and white calibration procedure should be performed before the initial use of the system. Besides, the whole procedure of spectral acquisition was performed in a dark room to avoid the interference of stray light.

A digital refractometer with a resolution of 0.1% Brix (PAL-1, ATAGO Inc., Japan) was utilized to measure the SSC of kiwifruit samples. First, the prepared kiwifruit juice was carefully dropped into the detection tank. Then, the SSC physicochemical values of SSC were recorded once the display data stabilized. It is worth noting that before measuring the SSC of each sample, it is essential to calibrate the refractometer reading by setting it to zero using distilled water. This calibration step was crucial to ensure the accuracy and reliability of the SSC measurements by accounting for any potential offset or drift in the refractometer readings.

2.3 Feature engineering

Feature engineering involves two key aspects: spectral pre-processing methods and spectral feature selection methods. Spectral pre-processing refines spectral reflectance by mitigating disturbances, while spectral feature selection eliminates redundancy, pinpointing crucial informative attributes for modeling. This duality is essential for extracting meaningful patterns from raw data and is imperative for developing robust regression models.

Recognizing that feature quality significantly influences model success, we implement an orthogonal experimental design for feature engineering. This methodical approach ensures experimentation and validation tailored to our specific modeling

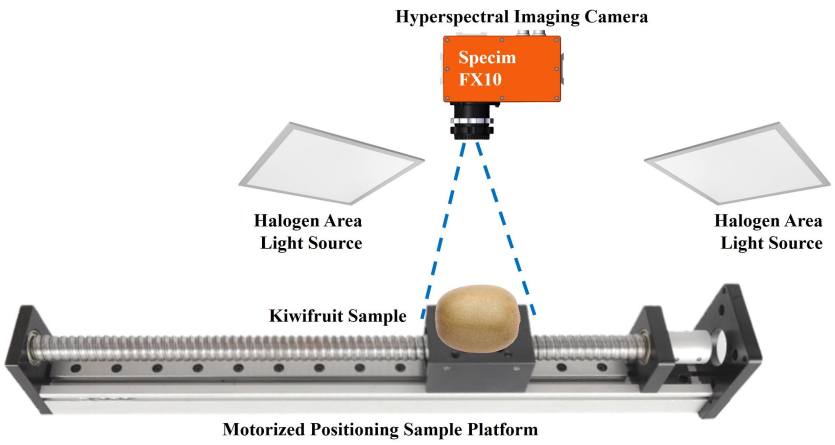


FIGURE 1
The custom-built hyperspectral imaging system.

context, enabling a structured assessment of diverse feature engineering strategies' effects on model accuracy. We rigorously investigate four spectral pre-processing and nine spectral feature selection techniques, assessing their individual and combined effects. The ensuing sections, 2.4 and 2.5, will delineate these techniques, underscoring their roles in data refinement and feature optimization, ultimately contributing to the improved accuracy of our model.

2.4 Spectral pre-processing methods

During the spectral acquisition step, various disturbances, such as sample differences, environmental noise, and baseline drift, can affect the final captured spectral image (Xu et al., 2023). To mitigate these variations in spectral reflectance and emphasize the features related to SSC, a spectral pre-processing procedure is conducted. It is a critical step in feature engineering (Lee et al., 2022) and primarily aims to refine and cleanse the data by removing unwanted noise, correcting baseline drift, and addressing other disturbances. To tackle the specific variations encountered in spectral pre-processing, a wide array of algorithms has been developed, each possessing unique characteristics and catering to various aspects of the process. In the following content, we will provide a brief description of several widely used spectral pre-processing methods that will be utilized in this study later.

Firstly, the Standard Normal Variant transform (SNV) (Dong et al., 2022; Liu et al., 2022) is a notable method that is meticulously designed to address the detrimental effects of scattering and concentration-related influences. It achieves this by normalizing spectral reflectance across the entire wavelength range, effectively mitigating deviations, and nullifying the impact of extraneous factors. Secondly, the Direct Orthogonal Signal Correction (DOSC) (Westerhuis et al., 2001) method disentangles spectral reflectance into correlated and uncorrelated components. By leveraging the principles of multivariate statistics, it discriminates between valuable signal information and intrusive background perturbations. In addition, the Detrend Correction (DC) (Ai et al., 2022) method adeptly attenuates the disruptive interference of external noise. It accomplishes this by subtracting the trend-fitting lines, enabling a refined and noise-free characterization of intrinsic spectral attributes. Lastly, the Savitzky-Golay (SG) (Savitzky and Golay, 1964) convolution smoothing method emerges as an exemplary technique for spectral refinement. By utilizing weighted polynomial regression within moving windows, it effectively suppresses high-frequency noise while preserving essential spectral features.

2.5 Spectral feature selection methods

Spectroscopy instruments typically exhibit highly correlated spectral responses, particularly in adjacent wavelength regions, leading to redundant data. Additionally, not all wavelengths are relevant to the problem at hand, potentially impacting the accuracy and precision of results. Therefore, discriminative feature selection becomes critical to enhance model performance. A range of spectral feature selection methods was investigated to address these issues, which are integral to feature engineering. These methods aim to identify and retain informative features, reduce the feature space, improve computational efficiency, and prevent multicollinearity and overfitting. Nine distinct spectral feature selection methods were identified and classified into three categories: basis-vectors-based, statistical-measures-based, and iterations-based methods. Each category offers unique approaches to feature selection and is briefly described below.

2.5.1 Based on basis vectors

Dimensionality reduction techniques such as Principal Component Analysis (PCA) (Pearson, 1901; Hotelling, 1933) and Singular Value Decomposition (SVD) (Smithies, 1938) use linear combinations of basis vectors to simplify high-dimensional data. PCA prioritizes components based on explained variance, while SVD utilizes singular values. Additionally, Kernel Principal Component Analysis (KPCA) (Schölkopf et al., 1997) extends PCA by capturing nonlinear patterns through a higher-dimensional kernel-based feature space, providing greater flexibility in representing high-dimensional data and extracting nonlinear features. By selecting a subset of basis vectors and transforming, these dimensionality reduction methods effectively reduce the dimensionality of the data while endeavoring to preserve as much information as possible.

2.5.2 Based on statistical measures

Individual wavelength features can also be evaluated using statistical measures. The F-test assesses the significance of feature differences between classes. Features with high F-values indicate greater relevance. Thus, one can rank the features based on their F-values and select the top n features for further analysis or dimensionality reduction. Similarly, the Pearson product-moment correlation coefficient (PPMCC) measures linear correlations, while Mutual Information (MI) detects both linear and non-linear dependencies.

2.5.3 Based on iterations

Iterative feature selection methods systematically search the feature space to identify the most relevant features for a specific problem. These methods, through a process of selection and elimination, adaptively integrate criteria, performance metrics, or domain knowledge. The Recursive Feature Elimination (RFE) (Araújo et al., 2001) is one such method that employs a backward elimination technique to prune irrelevant features from a regression model. Starting with all features, RFE trains the model, ranks features by their impact on model performance, and iteratively discards the weakest until a targeted feature set size or stopping condition is reached. The Successive Projection Algorithm (SPA) (Soares et al., 2013) selects features by projecting data onto orthogonal hyperplanes, treating spectral feature selection as a constrained combinatorial optimization problem. SPA minimizes multicollinearity, thereby reducing redundancy and addressing ill-conditioning by preventing the propagation of superfluous features during calibration. The Competitive Adaptive Reweighted Sampling (CARS) (Li et al., 2009; Zhang et al., 2019) focuses on

discarding features with minor regression coefficients in the PLSR model, using adaptive reweighting and cross-validation to fine-tune feature selection. CARS' adaptability allows it to dynamically capture dataset characteristics, which may result in varying feature selections across iterations.

2.6 Experiment settings

2.6.1 Sample division

The Sample Set Partitioning Based on Joint X-Y Distances (SPXY) (Wang et al., 2022) method was employed to divide the entire dataset of 116 kiwifruit samples into a calibration set and a prediction set, with a ratio of 3:1. This hold-out partitioning technique ensures a representative distribution of samples across both sets, allowing for the evaluation of model performance on unseen data.

Furthermore, the number of selected features in the spectral feature selection methods was determined using 5-fold cross-validation on the calibration set. This approach optimizes the feature selection process by iteratively evaluating the performance of different feature subsets across various subsets of the calibration set. By employing cross-validation, the optimal number of selected features is achieved while mitigating the risk of overfitting and ensuring the robustness of the model's performance.

2.6.2 Evaluation metrics

Three metrics, namely the Root Mean Square Error (RMSE), the Mean Absolute Percentage Error (MAPE), and the Residual Prediction Deviation (RPD) were employed to evaluate the impact of feature engineering on the regression model. These evaluation metrics are calculated using the following Equations 1–3.

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

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

$$RPD = \frac{SD}{RMSE} \quad (3)$$

where $\hat{y}_i$ is the predicted value of the i^{th} sample, y_i is the measured value of the i^{th} sample, and N is the total number of samples in the prediction set. Additionally, SD is the standard deviation of the measured value of the N samples. It is important to note that the metrics calculated for the validation set ($RMSE_v$, $MAPE_v$ and RPD_v) represent the mean values obtained from cross-validation. Conversely, the metrics calculated for the prediction set ($RMSE_p$, $MAPE_p$ and RPD_p) represent the mean values obtained from a single prediction. The details of the sample division and metrics calculation can be found in Figure 2.

2.6.3 The regression model

To comprehensively evaluate the effectiveness of feature engineering, a linear regression model based on OLS was established using an orthogonal experimental design. The OLS model, known for its ability to minimize the sum of squared residuals, is a widely-used regression method and a suitable choice for modeling the relationship between the input features and the SSC values. Its simplicity and interpretability make it a solid foundation for analyzing and comparing the effects of feature engineering on the regression model's performance. Meanwhile, those orthogonal experiments allow for a thorough examination of the individual effects of spectral pre-processing methods and spectral feature selection methods, as well as the exploration of potential interactions between them. By systematically varying and controlling these factors, researchers can gain valuable insights into the impact of different feature engineering techniques on the overall performance of the regression model.

3 Results and discussion

3.1 Distribution of the spectral reflectance and SSC

The distribution range of spectral reflectance in different wavelength regions were shown in Figure 3. Notably, the distribution range below 500 nm appears narrower, indicating lower variance and suggesting that this region contains less information. Conversely, the distribution range above 750 nm is broader, indicating higher variability in spectral reflectance within

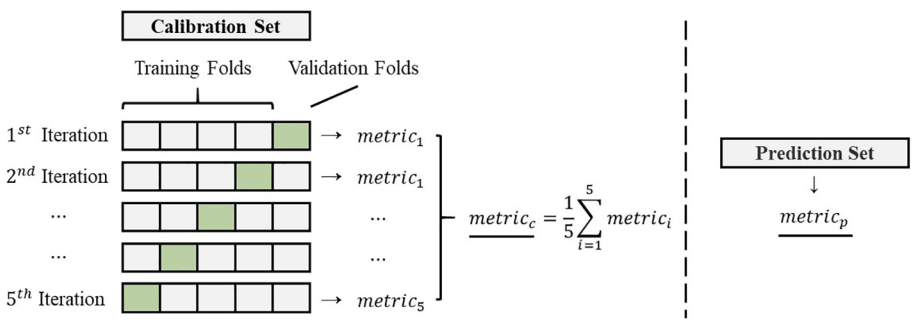


FIGURE 2
The details of the sample division and metrics calculation.

this wavelength region. This observation suggests that features of higher wavelength regions may contain more valuable information for the analysis and prediction of SSC values.

The SSC values for the complete dataset of 116 kiwifruit samples exhibit a mean value of 13.148 and a standard deviation of 1.025. The distribution of these values approximately follows a normal distribution, as evidenced by the Lilliefors test with a *p*-value of 0.0642. A visual representation of the frequency histogram depicting the specific distribution can be found in Figure 3. The calibration set of 87 samples presents a mean SSC of 13.165 and a standard deviation of 1.031, while the prediction set of 29 samples has a mean of 13.093 and a standard deviation of 1.022, indicating similar distribution parameters. Such comparability between calibration and prediction sets is vital to the reliability of our model's performance evaluation.

3.2 Regression performances

The performances of the OLS model under all conditions were summarized in Tables 1–5, grouped by spectral pre-processing methods, with the best scores highlighted in bold (due to rounding of specific metric values, some values that appear to be the same may still have minor differences). For a clearer comparison of outcomes among different spectral selection methods, we underline the results that fall below the baseline performance (i.e., without employing any spectral selection method) under identical spectral preprocessing conditions. The number of selected features of the corresponding spectral selection method is briefly represented by *n*.

These tables provide an exhaustive overview of the evaluation metrics, such as RMSE, MAPE, and RPD, enabling easy comparison and identification of the top-performing models within each feature preprocessing group. As shown in Tables 1–5, the superior performance of the OLS model utilizing feature engineering becomes evident when comparing it to the model without feature engineering. Within each spectral pre-processing method, employing a spectral feature selection method consistently enhanced performance across all metrics for both the calibration and validation sets (except for the minor anomaly of the $MAPE_p$ metric for the DC-CARS-OLS model).

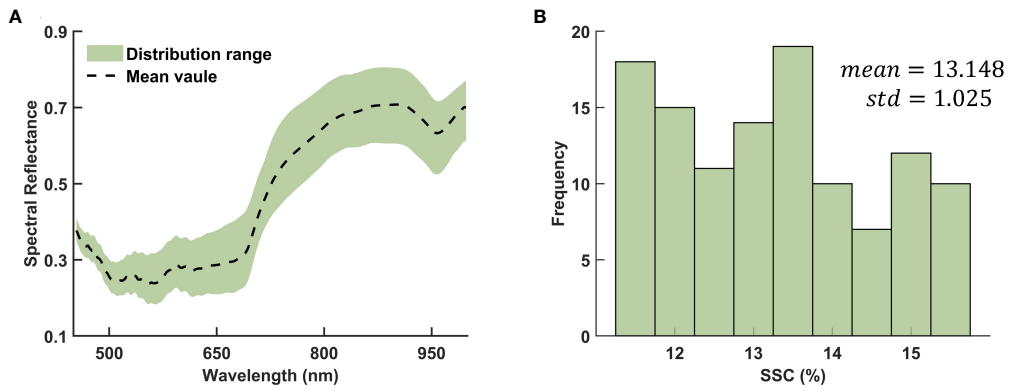


FIGURE 3
Distribution of the **(A)** spectral reflectance and **(B)** SSC.

TABLE 1 Regression performances using various spectral selection methods under no spectral pre-processing.

Pre-processing	Feature Selection	*n*	Calibration (Cross-Validation)			Validation		
			$RMSE_c$	$MAPE_c$	RPD_c	$RMSE_p$	$MAPE_p$	RPD_p
None	None	/	1.279	0.078	0.811	1.161	0.071	0.865
	PCA	5	0.953	0.062	1.061	0.780	0.053	1.288
	KPCA	5	0.947	0.062	1.068	0.780	0.053	1.287
	SVD	4	0.959	0.063	1.055	0.807	0.055	1.244
	F-test	10	0.966	0.061	1.050	0.943	0.059	1.064
	PPMCC	2	1.014	0.067	0.997	1.021	0.063	0.984
	MI	6	0.937	0.058	1.086	0.883	0.060	1.137
	RFE	6	0.994	0.065	1.017	0.773	0.050	1.299
	SPA	2	0.970	0.063	1.044	0.854	0.055	1.176
	CARS	8	**0.912**	**0.058**	**1.118**	**0.771**	**0.048**	**1.302**

The best scores are highlighted in bold.

TABLE 2 Regression results using various spectral selection methods under SNV spectral pre-processing.

Pre-processing	Feature Selection	*n*	Calibration (Cross-Validation)			Validation		
			$RMSE_c$	$MAPE_c$	RPD_c	$RMSE_p$	$MAPE_p$	RPD_p
SNV	None	/	1.795	0.109	0.593	1.535	0.098	0.654
	PCA	4	0.958	0.063	1.056	0.840	0.056	1.195
	KPCA	7	0.953	0.063	1.061	**0.740**	**0.048**	**1.358**
	SVD	4	0.958	0.063	1.056	0.840	0.056	1.195
	F-test	10	0.960	0.061	1.058	0.988	0.058	1.016
	PPMCC	2	1.018	0.067	0.995	1.030	0.064	0.975
	MI	4	1.001	0.064	1.019	0.948	0.061	1.059
	RFE	5	0.978	0.065	1.036	0.880	0.058	1.142
	SPA	8	**0.928**	**0.059**	**1.101**	0.795	0.049	1.263
	CARS	8	1.025	0.065	0.999	0.982	0.062	1.023

The best scores are highlighted in bold.

This conclusion, however, does not extend to spectral pre-processing methods. For the sake of simplicity, the performance outcomes of the OLS model under just a few selected spectral feature selection methods are succinctly summarized in Table 6. It is apparent that spectral pre-processing methods do not always lead to performance enhancements. Nevertheless, a judicious synergy between spectral pre-processing and feature selection methods may facilitate further amelioration of model performance. It is imperative for scholars to meticulously assess these variations when constructing an optimal feature engineering for their specific application.

These findings underscore the effectiveness of feature engineering in enhancing the regression model's predictive capabilities. In the calibration set, the DC-CARS-OLS model consistently demonstrates the best performance across all evaluation metrics ($RMSE_v$ = 0.760, $MAPE_v$ = 0.047 and RPD_v = 1.372), indicating that the combination of the DC spectral preprocessing method, the CARS spectral feature selection method, and the OLS regression model yields the most accurate and reliable predictions in this particular dataset. However, the performance differs in the validation set, where the SG-CARS-OLS model outperforms the other models, achieving the best scores in all evaluation metrics ($RMSE_p$ = 0.740, $MAPE_p$ = 0.046 and RPD_p = 1.358). This suggests that the combination of the SG spectral preprocessing method, the CARS spectral feature selection method, and the OLS regression model performs exceptionally well on unseen data. These findings emphasize the importance of evaluating model performance in both the calibration set and validation set to ensure the generalizability of the results.

TABLE 3 Regression results using various spectral selection methods under DOSC spectral pre-processing.

Pre-processing	Feature Selection	*n*	Calibration (Cross-Validation)			Validation		
			$RMSE_c$	$MAPE_c$	RPD_c	$RMSE_p$	$MAPE_p$	RPD_p
DOSC	None	/	/	/	/	/	/	/
	PCA	4	0.974	0.064	1.038	**0.809**	**0.052**	**1.242**
	KPCA	4	0.974	0.064	1.038	0.809	0.052	1.242
	SVD	3	0.984	0.065	1.029	0.845	0.055	1.188
	F-test	5	0.953	0.062	1.061	0.812	0.052	1.236
	PPMCC	5	0.953	0.062	1.061	0.812	0.052	1.236
	MI	2	0.980	0.064	1.033	0.864	0.054	1.162
	RFE	3	1.003	0.066	1.010	0.891	0.058	1.127
	SPA	13	0.970	0.061	1.053	0.879	0.057	1.142
	CARS	12	**0.888**	**0.053**	**1.173**	0.978	0.058	1.026

The best scores are highlighted in bold.

TABLE 4 Regression results using various spectral selection methods under DC spectral pre-processing.

Pre-processing	Feature Selection	*n*	Calibration (Cross-Validation)			Validation		
			$RMSE_c$	$MAPE_c$	RPD_c	$RMSE_p$	$MAPE_p$	RPD_p
DC	None	/	1.303	0.081	0.785	1.238	0.076	0.811
	PCA	31	0.934	0.058	1.119	0.802	**0.047**	1.252
	KPCA	9	0.957	0.060	1.069	**0.754**	0.051	**1.332**
	SVD	15	0.942	0.059	1.099	0.759	0.049	1.324
	F-test	1	0.964	0.063	1.050	0.831	0.053	1.208
	PPMCC	1	0.964	0.063	1.050	0.831	0.053	1.208
	MI	7	0.982	0.062	1.043	0.810	0.053	1.239
	RFE	5	1.004	0.066	1.011	0.810	0.053	1.240
	SPA	4	0.968	0.064	1.048	0.780	0.049	1.287
	CARS	18	**0.760**	**0.047**	**1.372**	1.189	0.077	0.844

The best scores are highlighted in bold.
The results that fall below the baseline performance are highlighted in underlined.

It further demonstrates that the optimal combination of feature preprocessing methods and spectral feature selection methods may vary depending on the dataset and the specific task concerned. Researchers should carefully consider these variations when designing the most suitable combination of feature engineering.

The frequency with which the OLS model achieves the best performance for each metric under every condition is summarized in Table 7. Among the spectral pre-processing methods, all exhibit an equal frequency of best performance. However, when considering spectral feature selection methods, it is noteworthy that the CARS method stands out with a significantly higher frequency of best performance compared to the other methods. This observation raises the possibility that greater attention should be directed toward spectral feature selection methods during the design of feature engineering and suggests that CARS is particularly effective in selecting informative features for enhancing the performance of the regression model.

3.3 Selected optimal features

The distribution of the features selected by the DC-CARS and SG-CARS methods are shown in Figure 4. The features extracted by the DC-CARS method show a more dispersed distribution across different wavelengths. In contrast, the features extracted by the SG-CARS method exhibit a relatively concentrated distribution, particularly around 600 nm and 850 nm. Both methods display a concentration of selected features above 750 nm, but there is also a smaller distribution near 600-700 nm. These findings align with the distribution range of spectral reflectance in different wavelength regions, as depicted in Figure 3A.

TABLE 5 Regression results using various spectral selection methods under SG spectral pre-processing.

Pre-processing	Feature Selection	*n*	Calibration (Cross-Validation)			Validation		
			$RMSE_c$	$MAPE_c$	RPD_c	$RMSE_p$	$MAPE_p$	RPD_p
SG	None	/	1.553	0.096	0.662	1.425	0.091	0.704
	PCA	5	0.953	0.062	1.061	0.780	0.053	1.287
	KPCA	5	0.948	0.062	1.068	0.780	0.053	1.287
	SVD	4	0.959	0.063	1.055	0.807	0.055	1.244
	F-test	8	0.989	0.063	1.029	0.924	0.059	1.087
	PPMCC	5	1.015	0.065	1.006	0.903	0.057	1.112
	MI	12	0.980	0.062	1.037	0.793	0.053	1.267
	RFE	6	0.963	0.063	1.049	0.774	0.052	1.298
	SPA	2	0.970	0.063	1.044	0.853	0.055	1.178
	CARS	13	**0.895**	**0.053**	**1.139**	**0.740**	**0.046**	**1.358**

The best scores are highlighted in bold.

TABLE 6 Regression results using various spectral pre-processing methods under no and CARS spectral pre-processing.

Pre-processing	Feature Selection	*n*	Calibration (Cross-Validation)			Validation		
			$RMSE_c$	$MAPE_c$	RPD_c	$RMSE_p$	$MAPE_p$	RPD_p
None	None	/	**1.279**	**0.078**	**0.811**	**1.161**	**0.071**	**0.865**
SNV		/	1.795	0.109	0.593	1.535	0.098	0.654
DOSC		/	/	/	/	/	/	/
DC		/	1.303	0.081	0.785	1.238	0.076	0.811
SG		/	1.553	0.096	0.662	1.425	0.091	0.704
None	CARS	8	0.912	0.058	1.118	0.771	0.048	1.302
SNV		8	1.025	0.065	0.999	0.982	0.062	1.023
DOSC		12	0.888	0.053	1.173	0.978	0.058	1.026
DC		18	**0.760**	**0.047**	**1.372**	1.189	0.077	0.844
SG		13	0.895	0.053	1.139	**0.740**	**0.046**	**1.358**

The best scores are highlighted in bold.
The results that fall below the baseline performance are highlighted in underlined.

3.4 The stacking generalization

We observed that the DC-CARS-OLS model, despite achieving the best performance in the calibration set, did not perform as well in the validation set. This suggests that the DC-CARS-OLS model may have overfit the calibration set and may not generalize well to unseen data. Conversely, the SG-CARS-OLS model achieved the best performance in the validation set but performed lower than the DC-CARS-OLS model in the calibration set, indicating that there is still room for improvement in its fitting ability.

TABLE 7 Statistics of the frequency of best performance for each metric under every condition.

	None	PCA	KPCA	SVD	F-test	PPMCC	MI	RFE	SPA	CARS	SUM
None	0	0	0	0	0	0	0	0	0	6	**6**
SNV	0	0	3	0	0	0	0	0	3	0	**6**
DOSC	0	3	0	0	0	0	0	0	0	3	**6**
DC	0	1	2	0	0	0	0	0	0	3	**6**
SG	0	0	0	0	0	0	0	0	0	6	**6**
SUM	0	4	5	0	0	0	0	0	3	**18**	/

The best scores are highlighted in bold.

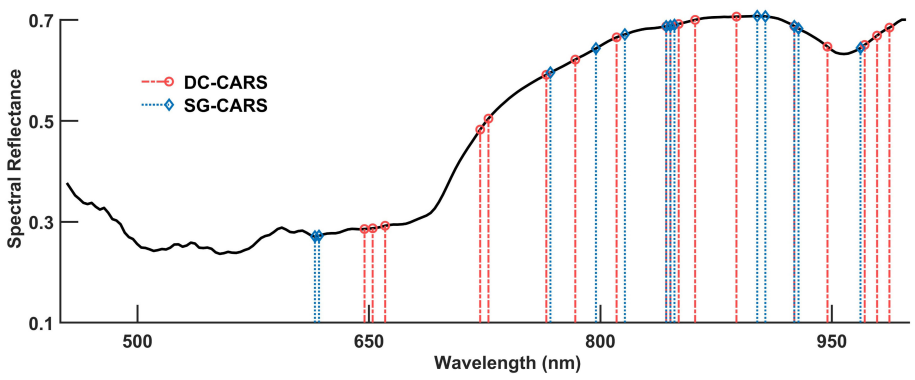


FIGURE 4
Distribution of the features selected by the DC-CARS and SG-CARS methods.

To leverage the strengths of both models and address these limitations, we introduced the stacking generalization technique (Wolpert, 1992). The stacking generalization technique is a powerful method that combines outputs of multiple base models to improve the final predictive performance. It involves constructing a meta-model that takes the predictions of base models as input, thus addressing the limitations of individual models and harnessing their complementary strengths. Specifically, the base models are trained on the same calibration dataset but with different methods or settings. The meta-model then learns to combine the outputs of base models in an optimal way to produce the final prediction. In this study, we utilized stacking generalization technique to combine the outputs of the DC-CARS-OLS model and SG-CARS-OLS model, aiming to leverage their respective strengths and enhance the final predictive capability and generalization performance of the regression model. The specific structure and computational flow of the stacking generalization model utilized in this study can be found in Figure 5, providing a visual representation of how the stacking generalization process is implemented.

The performance of the base models as well as the stacking generalization model is presented in Table 8. The performance of the stacking generalization model on the calibration set showed a decrease compared to the DC-CARS-OLS model. Besides, its performance has improved compared to the SG-CARS-OLS model on both the calibration and validation sets. These findings suggest that the stacking generalization model effectively addresses the overfitting issue observed in the DC-CARS-OLS model and further enhances the model's fitting ability based on the SG-CARS-OLS model. By combining the strengths of both base models, the stacking technique successfully achieves improved overall performance and enhanced generalization ability.

The comparison between the experimentally measured and stacking generalization model-predicted values of SSC is shown in Figure 6. The close alignment of predicted SSC distributions across both calibration and prediction datasets underscores the model's robustness, reflecting its capability to generalize well without succumbing to overfitting within the calibration phase.

This study's approach is benchmarked against established methods in the field, with comparative results detailed in Table 9. Moen et al. (Moen et al., 2021) explored the link between kiwifruit spectral data and SSC using various machine learning approaches and determined that the optimal prediction was achieved by UVE-PLS model, yielding an $RMSE_p$ of 1.047. Zhou et al. (Zhou, 2022) also investigated this relationship and discovered that SVR model offered the best predictive accuracy with an $RMSE_p$ of 1.309. Meanwhile, Benelli et al. (Benelli et al., 2022) applied a PLS

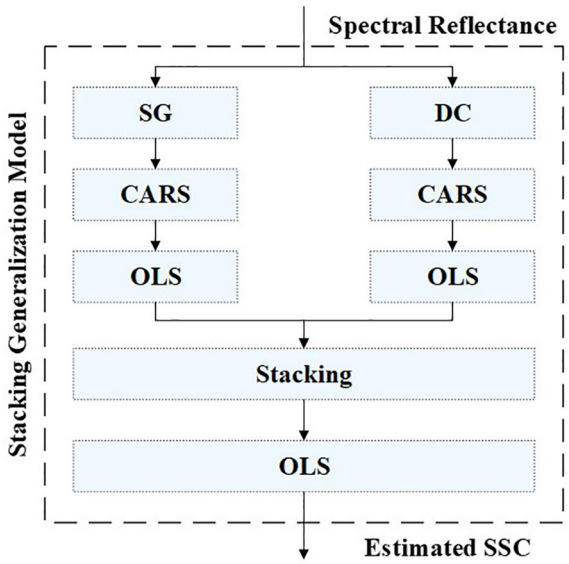


FIGURE 5
The specific structure and computational flow of the stacking generalization model.

TABLE 8 Regression results of the base models as well as the stacking generalization model.

Regression model	Calibration (Cross-Validation)			Validation		
	$RMSE_c$	$MAPE_c$	RPD_c	$RMSE_p$	$MAPE_p$	RPD_p
DC-CARS-OLS	**0.760**	0.047	1.372	1.189	0.077	0.844
SG-CARS-OLS	0.895	0.053	1.139	0.740	0.046	1.358
Stacking Generalization	0.782	**0.047**	**1.331**	**0.721**	**0.046**	**1.394**

The best scores are highlighted in bold.

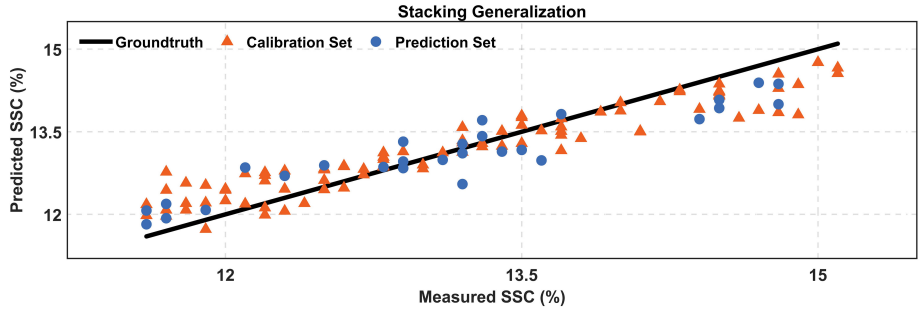


FIGURE 6
Comparison of the measured and predicted SSC.

TABLE 9 Comparison of the prediction results with the other methods.

Literature	Method	Calibration (Cross-Validation)	Validation
		$RMSE_c$	$RMSE_p$
Moen et al., 2021	UVE-PLS	/	1.047
Zhou, 2022	SVR	/	1.309
Benelli et al., 2022	PLS	0.810	0.730
This study	Stacking generalization	**0.782**	**0.721**

The best scores are highlighted in bold.

model leveraging hyperspectral imaging to assess kiwifruit maturity, attaining $RMSE_c$ and $RMSE_p$ values of 0.81 and 0.73, respectively. In our research, cross-validation within the calibration set was utilized to robustly detect overfitting, resulting in the most accurate predictions characterized by the lowest $RMSE_p$ in the validation set.

4 Conclusion

In conclusion, our investigation reveals that feature engineering, particularly the application of the CARS method for feature selection, significantly enhances SSC prediction accuracy in kiwifruit using hyperspectral imaging. Through rigorous comparative analysis, we established that the DC-CARS-OLS model delivers the most accurate results in calibration, while the SG-CARS-OLS model excels in validation scenarios. These outcomes specifically highlight the critical nature of spectral feature selection in constructing effective predictive models. Additionally, the introduction of the stacking generalization technique has proven instrumental in amalgamating the predictive strengths of individual models, thereby mitigating overfitting, and refining overall regression accuracy. Our findings not only bolster the methodological framework for non-destructive SSC estimation in kiwifruit but also suggest a template for broader application in agricultural quality assessment. The practical upshot of our study is a robust, non-invasive approach that promotes the kiwifruit industry's capability to ensure product quality, optimize resource use, and minimize waste. Ultimately, this research underlines the transformative potential of targeted feature engineering and advanced ensemble techniques in enhancing the precision of agricultural produce quality prediction models.

Data availability statement

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

Author contributions

WX: Data curation, Formal analysis, Investigation, Methodology, Writing – original draft. LW: Conceptualization, Investigation, Writing – review & editing. WC: Data curation, Writing – review & editing. XY: Visualization, Writing – review & editing. YL: Supervision, Writing – review & editing.

Funding

The author(s) declare financial support was received for the research, authorship, and/or publication of this article. This research is founded by National Key R&D Program of China (Project No.2021YFB3202500).

Acknowledgments

We express our sincere gratitude to the laboratory of Fulllight for generously providing the necessary instrumentation support, which was instrumental in the successful execution of this experiment. Additionally, we extend our heartfelt thanks to Licai

Xiao and Jun Song for their invaluable guidance and assistance in conducting the measurements.

Conflict of interest

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

Publisher's note

References

Ai, W., Liu, S., Liao, H., Du, J., Cai, Y., Liao, C., et al. (2022). Application of hyperspectral imaging technology in the rapid identification of microplastics in farmland soil. *Sci. Total Environ.* 807, 151030. doi: 10.1016/j.scitotenv.2021.151030

Araújo, M. C. U., Saldanha, T. C. B., Galvão, R. K. H., Yoneyama, T., Chame, H. C., and Visani, V. (2001). The successive projections algorithm for variable selection in spectroscopic multicomponent analysis. *Chemom. Intell. Lab. Syst.* 57, 65–73. doi: 10.1016/S0169-7439(01)00119-8

Benelli, A., Cevoli, C., Fabbri, A., and Ragni, L. (2022). Ripeness evaluation of kiwifruit by hyperspectral imaging. *Biosyst. Eng.* 223, 42–52. doi: 10.1016/j.biosystemseng.2021.08.009

Dong, C., An, T., Yang, M., Yang, C., Liu, Z., Li, Y., et al. (2022). Quantitative prediction and visual detection of the moisture content of withering leaves in black tea (Camellia sinensis) with hyperspectral image. *Infrared Phys. Technol.* 123, 104118. doi: 10.1016/j.infrared.2022.104118

Guo, Z., Chen, Q., Wu, J., Ouyang, Q., Chen, H., and Zhao, J. (2017). *Development of on-line detection system for simultaneous assessment of edible quality and internal defect in apple by NIR transmittance spectroscopy. in ASABE Paper No. 1701306* Vol. 1 (St. Joseph, MI: ASABE). doi: 10.13031/aim.201701306

Hotelling, H. (1933). Analysis of a complex of statistical variables into principal components. *J. Educ. Psychol.* 24, 417–441. doi: 10.1037/h0071325

Huang, Y., Lu, R., and Chen, K. (2018). Prediction of firmness parameters of tomatoes by portable visible and near-infrared spectroscopy. *J. Food Eng.* 222, 185–198. doi: 10.1016/j.jfoodeng.2017.11.030

Lee, A., Shim, J., Kim, B., Lee, H., and Lim, J. (2022). Non-destructive prediction of soluble solid contents in Fuji apples using visible near-infrared spectroscopy and various statistical methods. *J. Food Eng.* 321, 110945. doi: 10.1016/j.jfoodeng.2022.110945

Li, H., Liang, Y., Xu, Q., and Cao, D. (2009). Key wavelengths screening using competitive adaptive reweighted sampling method for multivariate calibration. *Anal. Chim. Acta* 648, 77–84. doi: 10.1016/j.aca.2009.06.046

Liu, Y., Long, Y., Liu, H., Lan, Y., Long, T., Kuang, R., et al. (2022). Polysaccharide prediction in Ganoderma lucidum fruiting body by hyperspectral imaging. *Food Chem. X* 13, 100199. doi: 10.1016/j.fochx.2021.100199

Ma, T., Li, X., Inagaki, T., Yang, H., and Tsuchikawa, S. (2018). Noncontact evaluation of soluble solids content in apples by near-infrared hyperspectral imaging. *J. Food Eng.* 224, 53–61. doi: 10.1016/j.jfoodeng.2017.12.028

Moen, J. E., Nilsen, V., Saidi, K. B., Kohmann, E., Devassy, B. M., and George, S. (2021). Hyperspectral imaging and machine learning for the prediction of SSC in kiwi fruits. *Nor. IKT-Konf. Forsk. Og Utdanning* 86–98(1), 86–98..

Nicolaï, B. M., Beullens, K., Bobelyn, E., Peirs, A., Saeys, W., Theron, K. I., et al. (2007). Nondestructive measurement of fruit and vegetable quality by means of NIR spectroscopy: A review. *Postharvest Biol. Technol.* 46, 99–118. doi: 10.1016/j.postharvbio.2007.06.024

Pearson, K. (1901). On lines and planes of closest fit to systems of points in space. *Lond. Edinb. Dublin Philos. Mag. J. Sci.* 2, 559–572. doi: 10.1080/14786440109462720

Pullanagari, R. R., and Li, M. (2021). Uncertainty assessment for firmness and total soluble solids of sweet cherries using hyperspectral imaging and multivariate statistics. *J. Food Eng.* 289, 110177. doi: 10.1016/j.jfoodeng.2020.110177

Savitzky, A., and Golay, M. J. E. (1964). Smoothing and differentiation of data by simplified least squares procedures. *Anal. Chem.* 36, 1627–1639. doi: 10.1021/ac60214a047

Schölkopf, B., Smola, A., and Müller, K.-R. (1997). "Kernel principal component analysis," in *Artificial Neural Networks — ICANN'97*. Eds. W. Gerstner, A. Germond, M. Hasler and J.-D. Nicoud (Berlin, Heidelberg: Springer), 583–588. doi: 10.1007/BFb0020217

Smithies, F. (1938). The eigen-values and singular values of integral equations. *Proc. Lond. Math. Soc* s2-43, 255–279. doi: 10.1112/plms/s2-43.4.255

Soares, S. F. C., Gomes, A. A., Araujo, M. C. U., Filho, A. R. G., and Galvão, R. K. H. (2013). The successive projections algorithm. *TrAC Trends Anal. Chem.* 42, 84–98. doi: 10.1016/j.trac.2012.09.006

Tian, S., Tian, H., Yang, Q., and Xu, H. (2022). Internal quality assessment of kiwifruit by bulk optical properties and online transmission spectra. *Food Control* 141, 109191. doi: 10.1016/j.foodcont.2022.109191

Wang, X., Xu, L., Chen, H., Zou, Z., Huang, P., and Xin, B. (2022). Non-destructive detection of pH value of kiwifruit based on hyperspectral fluorescence imaging technology. *Agriculture* 12, 208. doi: 10.3390/agriculture12020208

Westerhuis, J. A., de Jong, S., and Smilde, A. K. (2001). Direct orthogonal signal correction. *Chemom. Intell. Lab. Syst.* 56, 13–25. doi: 10.1016/S0169-7439(01)00102-2

Wolpert, D. H. (1992). Stacked generalization. *Neural Netw.* 5, 241–259. doi: 10.1016/S0893-6080(05)80023-1

Xiaobo, Z., Jiewen, Z., Povey, M. J. W., Holmes, M., and Hanpin, M. (2010). Variables selection methods in near-infrared spectroscopy. *Anal. Chim. Acta* 667, 14–32. doi: 10.1016/j.aca.2010.03.048

Xu, L., Chen, Y., Wang, X., Chen, H., Tang, Z., Shi, X., et al. (2023). Non-destructive detection of kiwifruit soluble solid content based on hyperspectral and fluorescence spectral imaging. *Front. Plant Sci.* 13. doi: 10.3389/fpls.2022.1075929

Yao, Y., Chen, H., Xie, L., and Rao, X. (2013). Assessing the temperature influence on the soluble solids content of watermelon juice as measured by visible and near-infrared spectroscopy and chemometrics. *J. Food Eng.* 119, 22–27. doi: 10.1016/j.jfoodeng.2013.04.033

Zhang, D., Xu, Y., Huang, W., Tian, X., Xia, Y., Xu, L., et al. (2019). Nondestructive measurement of soluble solids content in apple using near infrared hyperspectral imaging coupled with wavelength selection algorithm. *Infrared Phys. Technol.* 98, 297–304. doi: 10.1016/j.infrared.2019.03.026

Zhou, Y. (2022). Handheld nondestructive detector assessing internal quality of kiwifruit based on multispectral technology. Northwest A&F University. doi: 10.27409/d.cnki.gxbnu.2022.000529

frontiers | Frontiers in Plant Science

TYPE Original Research
PUBLISHED 12 February 2024
DOI 10.3389/fpls.2024.1344143

OPEN ACCESS

EDITED BY
Xi Tian,
Beijing Academy of Agriculture and Forestry Sciences, China

REVIEWED BY
Liu Zhang,
China Agricultural University, China
Hengchang Zang,
Shandong University, China
Leiqing Pan,
Nanjing Agricultural University, China

*CORRESPONDENCE
Junying Han
hanjy@gsau.edu.cn

RECEIVED 25 November 2023
ACCEPTED 24 January 2024
PUBLISHED 12 February 2024

CITATION
Zhu D, Han J, Liu C, Zhang J and Qi Y (2024) Modeling of flaxseed protein, oil content, linoleic acid, and lignan content prediction based on hyperspectral imaging. *Front. Plant Sci.* 15:1344143. doi: 10.3389/fpls.2024.1344143

Modeling of flaxseed protein, oil content, linoleic acid, and lignan content prediction based on hyperspectral imaging

Dongyu Zhu[1], Junying Han[1*], Chengzhong Liu[1], Jianping Zhang[2] and Yanni Qi[2]

[1]College of Information Science and Technology, Gansu Agricultural University, Lanzhou, China, [2]Crop Research Institute, Gansu Academy of Agricultural Sciences, Lanzhou, China

Protein, oil content, linoleic acid, and lignan are several key indicators for evaluating the quality of flaxseed. In order to optimize the testing methods for flaxseed's nutritional quality and enhance the efficiency of screening high-quality flax germplasm resources, we selected 30 flaxseed species widely cultivated in Northwest China as the subjects of our study. Firstly, we gathered hyperspectral information regarding the seeds, along with data on protein, oil content, linoleic acid, and lignan, and utilized the SPXY algorithm to classify the sample set. Subsequently, the spectral data underwent seven distinct preprocessing methods, revealing that the PLSR model exhibited superior performance after being processed with the SG smoothing method. Feature wavelength extraction was carried out using the Successive Projections Algorithm (SPA) and the Competitive Adaptive Reweighted Sampling (CARS). Finally, four quantitative analysis models, namely Partial Least Squares Regression (PLSR), Support Vector Regression (SVR), Multiple Linear Regression (MLR), and Principal Component Regression (PCR), were individually established. Experimental results demonstrated that among all the models for predicting protein content, the SG-CARS-MLR model predicted the best, with and of 0.9563 and 0.9336, with the corresponding Root Mean Square Error Correction (RMSEC) and Root Mean Square Error Prediction (RMSEP) of 0.4892 and 0.5616, respectively. In the optimal prediction models for oil content, linoleic acid and lignan, the R_p^2 was 0.8565, 0.8028, 0.9343, and the RMSEP was 0.8682, 0.5404, 0.5384, respectively. The study results show that hyperspectral imaging technology has excellent potential for application in the detection of quality characteristics of flaxseed and provides a new option for the future non-destructive testing of the nutritional quality of flaxseed.

KEYWORDS

hyperspectral imaging, flaxseed, protein, oil content, linoleic acid, lignan

1 Introduction

Flax (Linum usitatissimum) occupies an important position in oil and fiber crops (Oomah, 2001). According to its application scope, it is divided into fiber, oil, and fiber oil three (Zhang et al., 2011). Flaxseed is rich in essential omega-3 fatty acids, α-linolenic acid, and linoleic acid is recognized as a major source of high-quality proteins, lignan, lipids, and dietary fiber (Katare et al., 2012; Goyal et al., 2014), has a positive effect on the human diet and health, and its processed products in the world have a wide range of demand, belonging to the typical functional crops.

Currently, protein content in flaxseed is primarily determined through chemical analytical methods, like Kjeldahl nitrogen determination (Mueller et al., 2010; Yao et al., 2022). This method first requires drying and grinding of the sample, adding chemical reagents and heating, followed by distillation, titration treatment with a standard hydrochloric acid solution, and finally, a comprehensive calculation of the protein content results based on the values obtained from each process. Other methods for determining oil content often involve organic solvent extraction, while the quantification of linoleic acid and lignan is typically carried out using high-performance liquid chromatography (Meng et al., 2001; Feng et al., 2016). These traditional biochemical determinations of flaxseed nutrient content must be operated by professionals to complete the handling and operation process, which is both complex and professional, not only time-consuming and labor-intensive but also destructive to the sample and incidentally produces chemical pollution. To enhance the efficiency of screening high-quality flax germplasm resources, it is imperative to identify an accurate, rapid, and non-destructive method for assessing protein, oil content, linoleic acid, and lignan content.

HSI technology simultaneously captures the target's spatial characteristics and spectral information, effectively combining image and spectral data (Xiang et al., 2022). The spectral properties of an object are closely related to its intrinsic physicochemical properties, and the differences in the composition and structure of substances result in the selective absorption and emission of photons of different wavelengths within the substance. Presently, HSI serves as a non-destructive and expeditious analytical tool across various domains, including medical diagnosis (Bjorgan and Randeberg, 2015), food industry (Ma et al., 2019), fruit damage and disease detection (Tian et al., 2020; Yadav et al., 2022; Jiang et al., 2023), and plant seed analysis (Zhu et al., 2019). HSI has proven to be an effective technique for non-destructive seed quality testing by many scholars. For instance, Tu et al. (Tu et al., 2022) used HSI to detect similar maize authenticity. Zou et al. (Zou et al., 2023) employed HSI to gauge peanut seed vigor. In addition, Yoo et al. (Yoosefzadeh-Najafabadi et al., 2021) used HSI for soybean yield prediction. Zhang et al. (Zhang et al., 2022) Used HSI to detect hybrid wheat seed purity. Lu et al. (Lu et al., 2022) ingeniously combined HSI with deep convolutional generative adversarial networks to predict the oil content of individual corn kernels. Yu et al. (Yu et al., 2016) effectively measured fat content in peanuts (R_p^2 = 0.84 and SEP = 1.88) and Ma et al. (Ma et al., 2021) further devised a streamlined model for the non-destructive assessment of protein content in rice, achieving notable success (R_p^2 = 0.8011 and RMSEP = 0.52). All of these studies demonstrated the feasibility of detecting seed quality based on HSI. However, few studies have been reported on HSI detection of the internal quality of flaxseed. Leomara Floriano Ribeiro et al. employed infrared reflectance spectroscopy and multivariate correction to predict linolenic and linoleic acid content in flaxseed, achieving prediction sets with R_p^2 values as high as 0.90 and 0.86 (Ribeiro et al., 2013). While this method achieves high accuracy, it is limited to determining the content of linolenic and linoleic acids in only two types of flaxseed. Currently, with over 5,000 flax varieties in commercial cultivation, each exhibiting significant variations in nutrient composition, the method lacks generalizability and stability, rendering it ineffective for the determination of other species. Party Zhao et al. used near infrared analysis technology to determine the quality of flax germplasm resources, and Ye Jiali et al. used non-destructive near infrared spectroscopy to quantitatively analyze the content of flax seed protein, linolenic acid, and lignan (Dang and Zhao, 2008; Ye et al., 2021). The above three non-destructive tests on the nutritional quality of flaxseed are used in the infrared spectrometer wavelength range of 1100-2500 nm, 900-1700 nm, and 1000-2499 nm. The wavelength range of the imaging instrument, although high precision, the cost is expensive; the processing and operation of the process are both complex and professional, and it is not only not applicable to field operations but also general scientific researchers and flax planting researchers cannot be realized. In addition, these methods might not completely capture the internal features of the specimen, and they are solely employed to acquire spectral details from a solitary point source. The uniformity of the sample distribution consistently influences this and may not be the optimal selection. (Ozaki, 2021; Hu et al., 2023).

This project is dedicated to studying the 400-1000nm spectral range of flaxseed nutritional quality detection to fill the existing band range of research gaps. The spectral range of imaging instruments is relatively common and inexpensive. General researchers and flax planting researchers can easily buy and use. This study simultaneously analyzed the flaxseed protein, oil content, linoleic acid, and lignans' 4 nutrient content. Common reports of up to 3 nutrients have been analyzed in the literature. From the results of the literature available from multiple sources, it is the first time that the content of four nutrients was analyzed simultaneously. Additionally, comprehensively detecting multiple indicators of flaxseed allows for a more integrated assessment of its quality. Various nutrients in flaxseed are interconnected; therefore, solely predicting a single nutritional indicator is insufficient for quality measurement. Practical significance is achieved only through a simultaneous and comprehensive evaluation of several indicators. This integrated research approach contributes to a more thorough, systematic understanding and utilization of the potential value of flaxseed. Thus, this study seeks to establish a non-destructive and expeditious method utilizing HSI for detecting protein content, oil content, linoleic acid, and lignan in flaxseed. The primary research objectives encompass: (1) establish a PLSR prediction model of flaxseed protein content based on raw and preprocessed spectra and determine the optimal preprocessing method through model evaluation; (2) construct prediction models for flaxseed protein, oil content, linoleic acid, and lignan

based on distinctive wavelengths extracted by SPA and CARS, using PLSR, PCR, SVR, and MLR. The selection of the optimal prediction model for flaxseed protein, oil content, linoleic acid, and lignan relies on R_p^2 and RMSEP values to achieve swift, non-destructive, and precise nutritional quality prediction; (3) identifying characteristic spectral bands pertinent to protein, oil content, linoleic acid, and lignan in flaxseed based on the most effective model.

2 Materials and methods

2.1 Experimental materials

As shown in Table 1, thirty flaxseed varieties, extensively cultivated in Northwest China, were selected for the study. Seed samples were obtained from the Gansu Academy of Agriculture's Crop Institute. All the varieties were harvested in 2022 from the experimental field of Lanzhou New District, Gansu Province, China, situated at an altitude of 1520 m above sea level (103° 72'E,36°03'N). To limit water absorption, the flaxseeds were stored in sealed paper bags. Every sampling session involved collecting fifty intact and undamaged flaxseeds from each variety. Following acquiring hyperspectral images, they were immediately dispatched to the Gansu Academy of Agricultural Sciences in China to analyze protein, oil content, linoleic acid, and lignan for each variety.

2.2 Hyperspectral image capture

2.2.1 Hyperspectral imaging system

The Gaia Field portable hyperspectral system (Sichuan Dualix Spectral Imaging Technology Co., Ltd) is shown in Figure 1, which includes GaiaField-V10E hyperspectral camera, 2048×2048 pixels imaging lens, HSI-CT-150×150 standard whiteboard (PTFE), HSIA-DB indoor imaging dark box, four groups of shadowless lamp light source, HSIA-TP-L-A tripod rocker set, and hyperspectral data acquisition software Spec View. The spectral range is 380-1018 nm, spectral bands are 320, spectral resolution is 2.8 nm, the numerical aperture is F/2.4, slit size is 30 μm× 14.2 μm, the detector is SCMOS, and the imaging mode is built-in push-scan, autofocus, and dynamic range is 14 bits. The core components of the hyperspectral equipment include a standardized light source, a spectral camera, an electronically controlled mobile platform, a computer, and control software. The working principle is that the system adopts the push-scan imaging mode, the surface array detector and the imaging spectrometer are combined, and under the drive of the scanning control electric moving platform, the slit of the imaging spectrometer and the focal plane of the imaging lens undergoes relative motion, the detector collects real-time information relative to the line target, and finally splices into a complete cube of data.

2.2.2 Image acquisition and calibration

Enact the hyperspectral instrument switch and the dark box light source before image acquisition. Allow a 30-minute warm-up period, then configure the instrument parameters, setting the camera exposure time to 49ms, gain to 2, frame rate to 18.0018Hz, and forward speed to 0.00643cm/s. We have selected a total of 30 varieties of flaxseed; for each variety of hyperspectral images were collected a total of three times, each time from the corresponding varieties of randomly selected 50 seeds placed in the dark box on the mobile platform, as shown in Figure 1, and then these 50 seeds as the same ROI, to get an average spectral curve of these 50 seeds. After one acquisition for each variety, the sample under test was re-poured into the sample bag and shaken manually. Then, 50 seeds were randomly taken out for the subsequent image acquisition of that variety, repeated three times to get three average spectral curves and a total of 150 seeds were scanned. Ninety acquisitions were made for 30 varieties, with 4,500 seeds scanned, and 90 average spectral curves were obtained. After completing the acquisition, the original hyperspectral images underwent black-and-white correction to eliminate dark current noise introduced

TABLE 1 Flaxseed varieties.

No.	Variety	No.	Variety	No.	Variety
1	Onyc	11	Hua Ya 5	21	Yi Ya 3
2	Shuang You Ma 1	12	Hua Ya 6	22	Ba Ya 18
3	Shuang Ya 12	13	Ding Ya 17	23	Ba 14
4	Shuang Ya 14	14	Hei Ya 2	24	901 Ba Ya 15
5	Shuang Ya 15	15	Ning Ya 10	25	139 Ba Ya 17
6	Zhang Ya 3	16	Ba 9	26	Hua Ya 1
7	Ba 6	17	Ba 11	27	Hua Ya 2
8	Ba 5	18	Gan Ya 3	28	Hua Ya 3
9	Ba 4	19	Yan Za 10	29	Hua Ya 4
10	Ba 3	20	Jin Ya 7	30	Ba 2

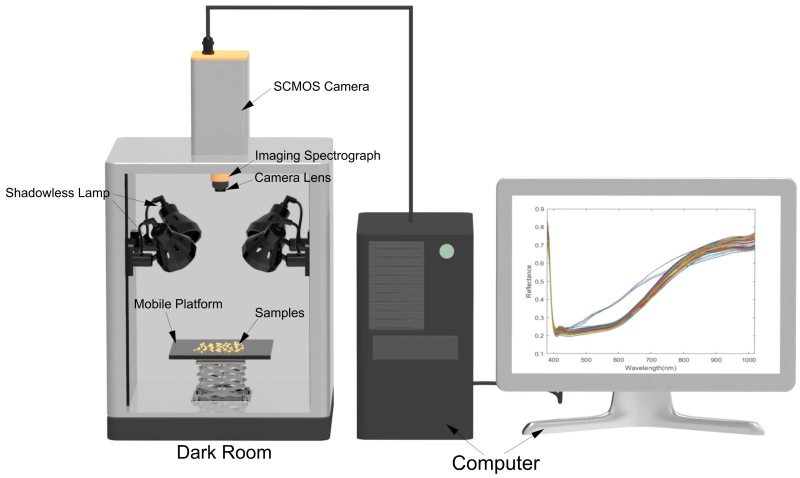


FIGURE 1
The hyperspectral imaging system.

by the camera. (Wang et al., 2022). The black-and-white correction formula is shown in Equation (1):

$$I_c = \frac{I_{raw} - I_{dark}}{I_{white} - I_{dark}} \quad (1)$$

Where I_{raw} is the raw image, I_{white} is the white reference image, I_{dark} is the dark reference image, and I_c is the calibrated image.

In order to extract the spectral information from the corrected hyperspectral image, the 50 flax seed region in a single image was used as the region of interest, and the spectral information was extracted, as shown in Figure 2. Firstly, the regions of interest (ROIs) of flax seeds and background were created separately in ENVI5.3 software, and then according to the different ROIs, the flax seeds and background were classified using support vector machine (SVM) in supervised classification and transformed into vectors, followed by masking process and transformed into mask images. Applying the mask image to the original hyperspectral image separates the hyperspectral image of all the flaxseed sample regions from the background to get the region of interest for the whole sample. Finally, it calculates the average of the spectra of all the flaxseeds on the hyperspectral image as the spectrum of that sample.

2.3 Sample Content Determination and Segmentation

The protein, oil content, linoleic acid, and lignan contents of 30 varieties of flaxseed were determined by the Gansu Academy of Agricultural Sciences in China. Sample set partitioning based on joint X - Y distances (SPXY) (Liu et al., 2011) was employed to allocate flaxseed protein, oil content, linoleic acid, and lignan into modeling and prediction sets at a 2:1 ratio. The reasonableness of the sample division was assessed by calculating the samples' maximum, minimum, average, and standard deviation in the

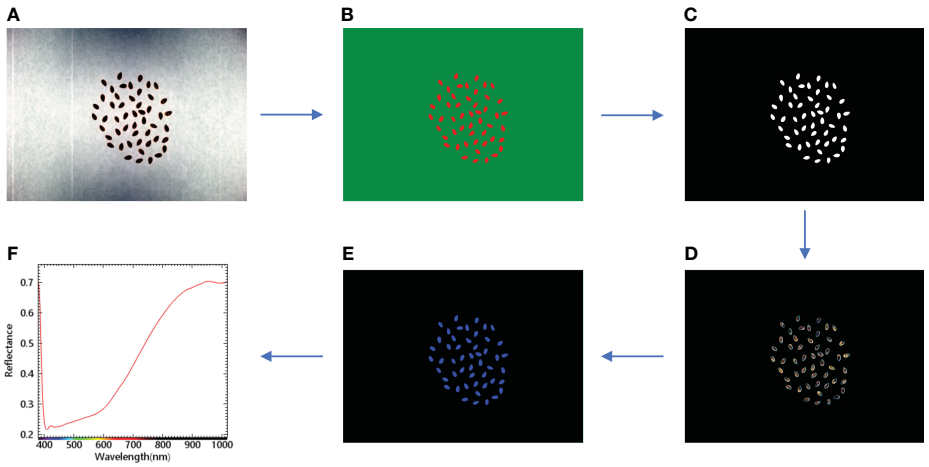


FIGURE 2
Sample hyperspectral image classification mask and spectral extraction flowchart. **(A)** Hyperspectral image; **(B)** Classification image; **(C)** Mask image; **(D)** Application mask image; **(E)** Region of interest image; **(F)** Average spectral curve.

training and prediction sets (Shao et al., 2020). The results are shown in Table 2. The maximum and minimum values of the training set for protein, oil content, and lignan included the prediction set, and the minimum values of the training set for linoleic acid and the prediction set were almost the same. Therefore, the overall division of the sample set is deemed reasonable.

2.4 Spectral preprocessing methods

During the acquisition of raw spectral data, it is often subject to various noise interferences, such as instrumental noise and environmental interference. In order to improve the quality and analyzability of the data, the extracted spectral information better reflects the changes in the sample curves to ensure that accurate and reliable results are obtained when building predictive models or conducting analyses. Therefore, it is necessary to pre-process the raw spectra to eliminate the noise as much as possible or reduce the influence of other environmental factors on the spectral information. The study employed various preprocessing techniques (Savitzky-Golay (SG) smoothing, normalization, baseline, standard normal variable correction (SNV), moving average (MA), multiple scattering correction (MSC), and first-order derivative (1st Der)) on the raw flaxseed spectra (Aulia et al., 2023). SG is mainly used to achieve the effect of smoothing curves and reducing noise by fitting local polynomials to the original spectra using a sliding window; Normalize can normalizes the spectral data to the same scale, which usually scales the value of each wavelength to a value between 0 and 1. It is mainly used to eliminate intensity differences due to differences in spectral measurement instruments, measurement conditions, and other factors; Baseline is based on the principle of removing baseline fluctuations in the spectrum due to instrumental drift, background changes, and other reasons, and can be used to improve the accuracy of the data; SNV is standardized by calculating the ratio of the spectral value at each wavelength to the mean and standard deviation of all spectral values at that wavelength; The aim is to reduce the intensity differences in the spectra and highlight the chemical information; MA focuses on averaging the spectral data over a sliding window to reduce high-frequency noise and smooth the spectral curves; MSC is based on the principle of correcting for multiple scattering by comparing the spectral data with a selected reference spectrum. This includes fitting each spectrum to the mean using least squares regression and calculating the preprocessed data by decomposing the slope and intercept of the regression. The aim is to reduce the effect of multiple scattering and emphasize the chemical information to improve the accuracy of quantitative analysis; 1st Der is to perform first-order derivative operations on the spectral data to highlight the rate of change of the spectral lines, enhance the peaks and valleys in the spectra, and highlight spectral line features. Subsequently, a PLSR prediction model for the protein content of flaxseed was established based on the raw and pretreatment spectra, and the optimal pretreatment method was determined by model evaluation.

2.5 Feature band extraction methods

Various sources frequently disrupt raw spectral data acquisition. Since the full spectrum contains 320 wavelength variables, not all wavelengths are useful for the analysis task. Extracting characteristic wavelengths reduces data dimensions, eliminates redundancy, and enhances modeling efficiency and performance. This study employs the successive projections algorithm (SPA) and the competitive adaptive reweighted sampling (CARS) algorithm for wavelength feature extraction. SPA algorithm is a forward looping feature variable selection method, which is a method of selecting feature wavelengths by calculating the correlation between each wavelength and the target variable, which is capable of filtering out the invalid information and greatly reducing the influence of covariance among the data. SPA has intuition and simplicity for the downscaling and feature selection of spectral data, which makes the model easier to interpret and understand (Li et al., 2023). CARS is an innovative variable selection algorithm proposed by Li (Li et al., 2009). At the same time, CARS is also a commonly used method for selecting the characteristic wavelengths, which firstly utilizes the PLS model to screen the wavelengths with large regression coefficients and then optimally selects the wavelengths with the smallest root-mean-square error through ten-fold cross-validation A subset of wavelengths is selected through ten-fold cross-validation, and the most critical variable for the prediction target is selected as the wavelength. The CARS algorithm is more flexible and adaptive than the traditional weighting methods, which helps to retain more useful information. In addition, CARS can more fully consider the correlation between wavelengths, thus better reflecting the

TABLE 2 Flaxseed protein, oil content, linoleic acid, and lignan sample set contents.

Sample set	Protein		Oil content		Linoleic acid		Lignan	
	Cal	Pre	Cal	Pre	Cal	Pre	Cal	Pre
Number of samples	60	30	60	30	60	30	60	30
Maximum (%)	28.46	27.76	40.9	40.5	13.81	13.58	11.06	8.39
Minimum (%)	23.01	23.07	33.38	34.65	9.93	9.92	4.79	5.67
Average (%)	25.1	25.21	36.5	36.4	11.96	11.93	8.14	7.38
Standard deviation	1.54	1.28	1.62	1.5	0.86	0.82	1.49	0.71

characteristics of the data. In hyperspectral data, the CARS algorithm helps select representative characteristic wavelengths more comprehensively, considering that there may be complex relationships between wavelengths (Xu et al., 2022).

2.6 Modeling methods

Partial least squares regression (PLSR) is a multivariate statistical method (Wang et al., 2019). PLSR models the spectral data by minimizing the covariance between the spectral data and the target variable. It achieves data downscaling by introducing latent variables and then regressing these latent variables on the target variables.

Support vector regression (SVR) can fit data quickly (Xiang et al., 2022), and it deals with nonlinear relationships by mapping the data into a high-dimensional space and then constructing a linear regression model in that space.

Principal component regression (PCR) models spectral data by downscaling them into principal components to explain the variance of the spectral data and predict the target variable (Mahesh et al., 2015).

Multiple linear regression (MLR) is a conventional linear regression method that establishes the relationship between multiple independent variables and the dependent variable. In MLR, each wavelength is treated as a predictor variable, and the model tries to find a linear combination between these variables to fit the target variable best. However, MLR modeling only applies when the number of variables is less than the number of samples. Consequently, in this study, only wavelengths extracted by CARS and SPA algorithms were used for modeling (Rajkumar et al., 2012).

2.7 Software and model assessment

Besides using Spec view software for hyperspectral image acquisition and ENVI 5.3 for spectrum extraction, we utilized 3ds Max to construct a 3D model of the HSI system. Unscrambler X handled spectrum preprocessing and model building, while MATLAB R2021b extracted the featured wavelengths and plotted the waveforms. This paper assesses the model's performance using various evaluation metrics, including the cross-validation correlation coefficient (R_{cv}^2) and root mean square error (RMSECV), the calibration set correlation coefficient (R_c^2) and root mean square error (RMSEC), and the prediction set correlation coefficient (R_p^2) and root mean square error (RMSEP) (Zhang and Guo, 2020). The calculation process is detailed in Equation (2) and Equation (3). A well-performing model is characterized by high R_{cv}^2, R_c^2, or R_p^2 values and low RMSECV, RMSEC, or RMSEP values. These metrics gauge the model's fitting and prediction capabilities, ensuring it excels in data fitting and new data prediction. The processing of the whole experiment is shown in Figure 3.

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

$$RMSE = \sqrt{\frac{1}{n}\sum_{i-1}^{n}(y_i - \hat{y}_i)^2} \tag{3}$$

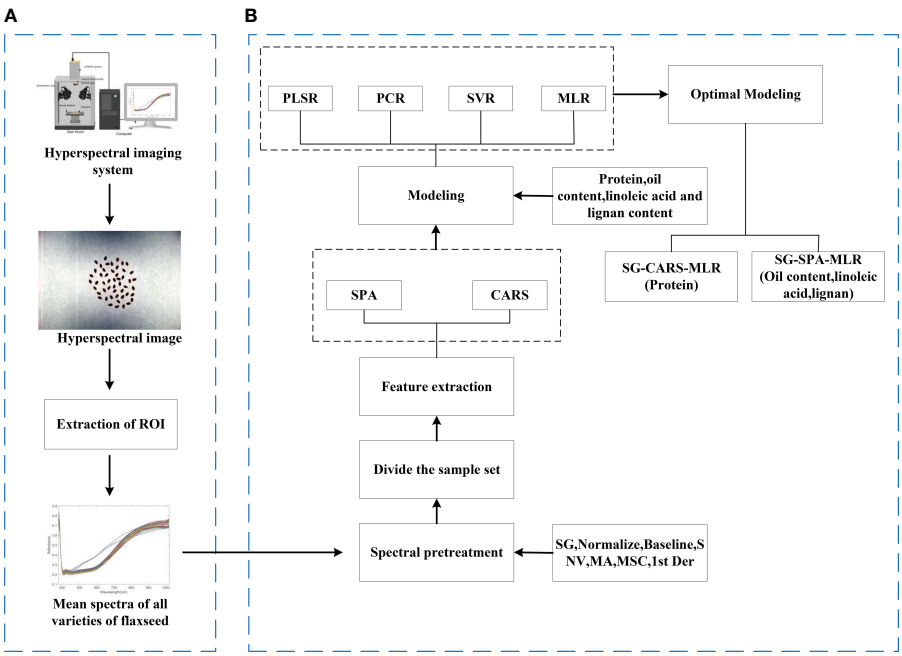


FIGURE 3
Experimental procedure. **(A)** Process of raw hyperspectral image acquisition and ROI extraction. **(B)** Spectral preprocessing, feature extraction, and modeling processes.

3 Results and analyses

3.1 Spectral characterization and selection of optimal preprocessing

Figure 4 shows the average spectra of 30 different flaxseed varieties and the average spectra of 7 pre-treatments containing a total of 4,500 samples. As evident from Figure 4A, the average spectral profiles of various flaxseed varieties exhibit a consistent trend. However, notable deviations appear in the 450-800nm range, likely attributable to variations between flaxseed varieties. Further studies revealed that the average spectral profile of flaxseed has a significant reflectance peak at 420 nm, which is mainly caused by carotenoids (Yang et al., 2021). In addition, the spectral profile shows a clear upward trend in the range of 600-750 nm, which is attributed to the fact that this wavelength corresponds to the vibration of the N-H chemical bond of amino acids in the seeds (Xu et al., 2022). The absorption peak near 980 nm originates from the O-H stretching vibration, which is related to the structure of water molecules (Yu et al., 2014).

To minimize the influence of noise and irrelevant information in spectral data, preprocessing of raw spectral information is essential. The Partial Least Squares Regression (PLSR) model

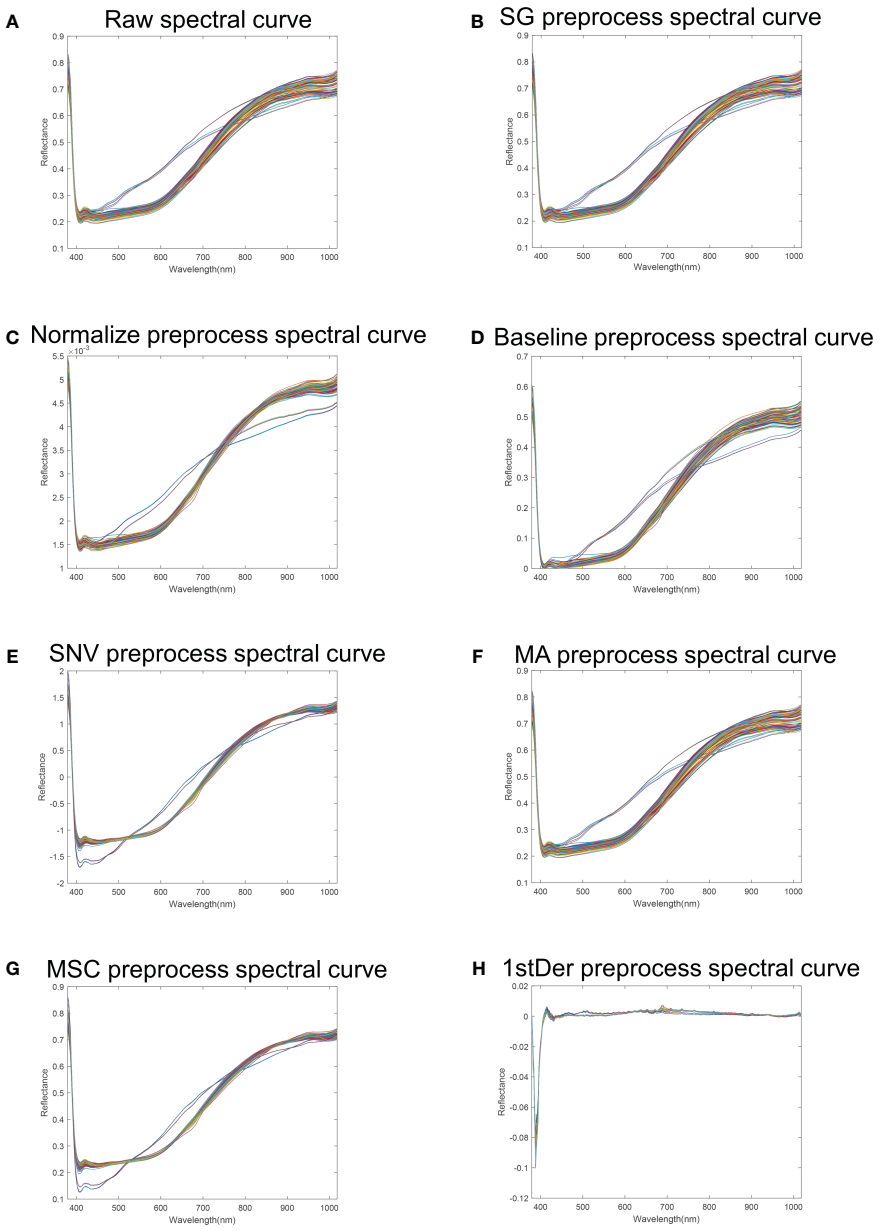


FIGURE 4
Flaxseed spectral reflectance curves. **(A)** Raw spectral curve of flaxseed; **(B)** SG preprocess spectral curve of flaxseed; **(C)** Normalize preprocess spectral curve of flaxseed; **(D)** Baseline preprocess spectral curve of flaxseed; **(E)** SNV preprocess spectral curve of flaxseed; **(F)** MA preprocess spectral curve of flaxseed; **(G)** MSC preprocess spectral curve; **(H)** 1stDer preprocess spectral curve.

comprehensively addresses the relationship between independent and dependent variables, even in scenarios of significant multicollinearity. The PLSR model for predicting flaxseed protein content identifies the best preprocessing method using stochastic cross-validation, employing Cross-validation set R^2_{cv} and RMSECV as model evaluation metrics. Figure 5 illustrates that, among the PLSR models predicting flaxseed protein content without pretreatment and with seven different pretreatment methods, the SG-PLSR model offered superior results, displaying a R^2_{cv} value of 0.8394 and an RMSECV value of 0.6010. Thus, the SG pretreatment method was adopted for further feature extraction in predicting oil content, linoleic acid, and lignan content.

3.2 Results of feature extraction

Figures 6A, B shows the wavelength distribution of flaxseed protein characteristics selected by the SPA algorithm, specifying the number of variables N = 1 to 30. When the variable is 14, the RMSE value is the smallest. Therefore, the final number of wavelengths selected is 14, accounting for 4.3% of the total number of wavelengths. These wavelengths, displayed in Figure 6B, correspond to the variables 391, 394, 405, 408, 424, 440, 465, 491, 640, 793, 842, 902, 990 nm and 1014 nm, respectively.

Figure 7 shows the process of selecting the characteristic wavelengths of flaxseed proteins by the CARS algorithm, which includes the relationship between the number of sampling runs and

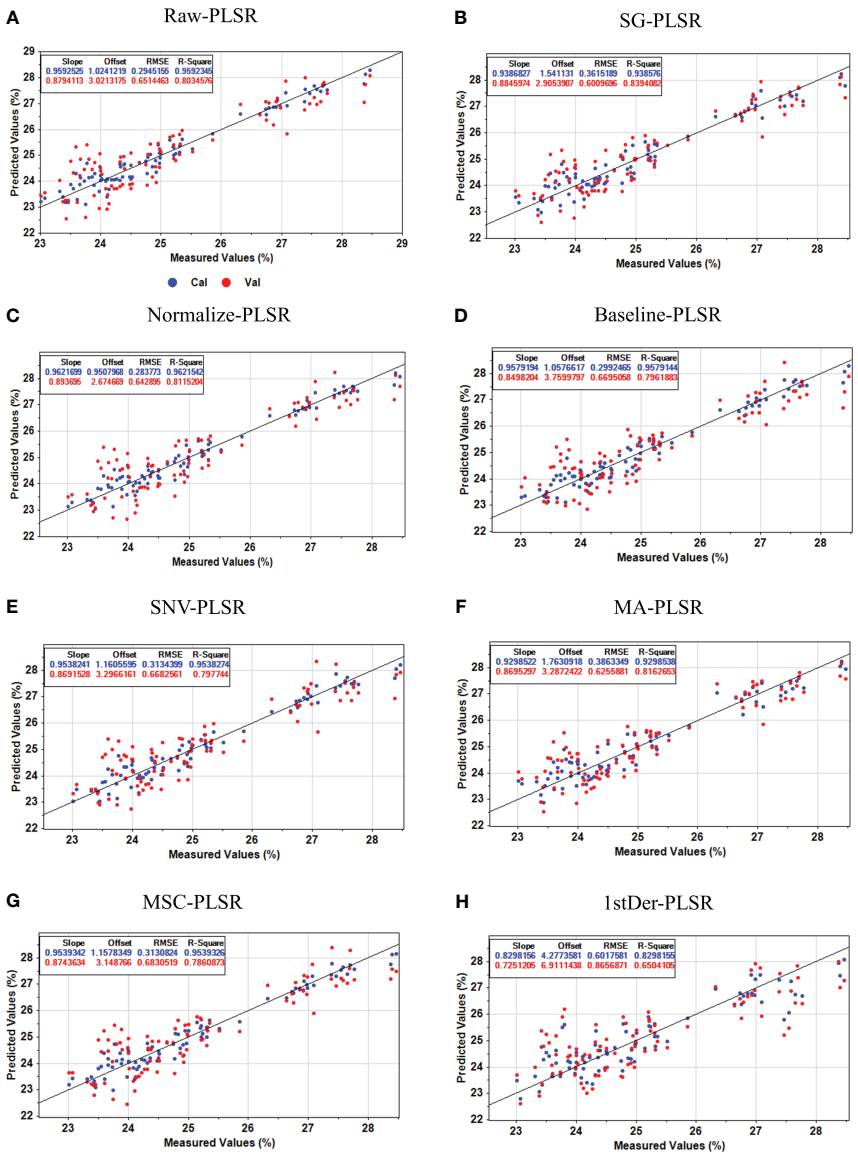


FIGURE 5
Protein content prediction results of the PLSR model based on different preprocesses. **(A)** Raw-PLSR; **(B)** SG-PLSR; **(C)** Normalize-PLSR; **(D)** Baseline-PLSR; **(E)** SNV-PLSR; **(F)** MA-PLSR; **(G)** MSC-PLSR; **(H)** 1stDer-PLSR.

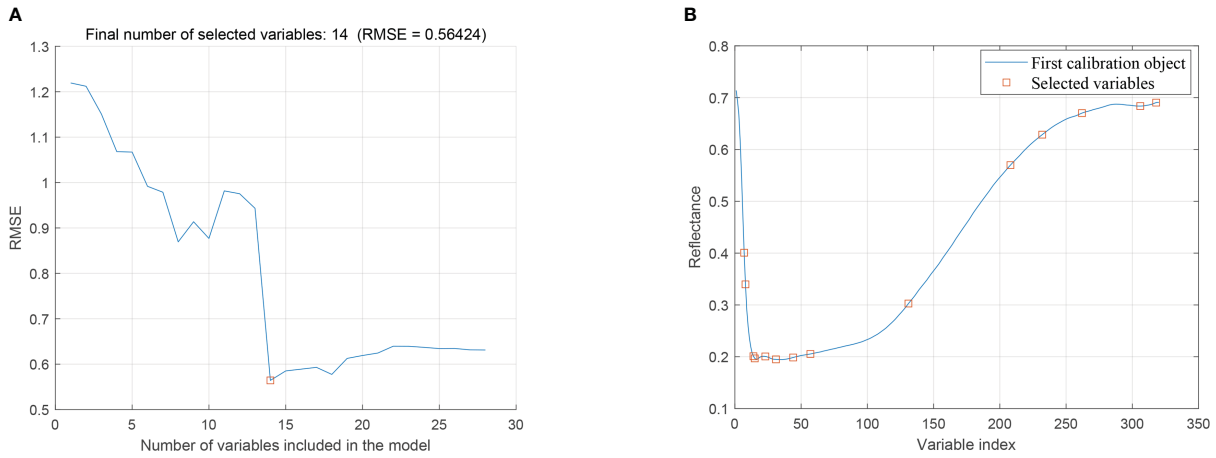


FIGURE 6
SPA extraction of feature variables. **(A)** Trend of RMSE with feature variables, **(B)** Distribution of preferred feature variables.

the number of selected wavelength variables, the relationship between the RMSECV values and the relationship between the regression coefficients path. This figure illustrates that the efficiency of feature variable selection significantly improves from rough to fine screening with the increased number of sampling runs. Moreover, when the number of runs reached 21, RMSECV minimized, selecting 33 characteristic wavelengths crucial for predicting protein content. These wavelengths include 405, 408, 424, 438, 441, 465, 468, 494, 497, 501, 517, 519, 529, 569, 571, 574, 576, 593, 595, 598, 772, 844, 846, 880, 910, 931, 933, 958, 960, 986, 988, 1009 nm and 1014 nm, amounting to 10.3% of the total wavelength. This process indicates removing substantial irrelevant hyperspectral data and flaxseed protein content prediction in runs 1 to 20. The SPA and CARS methods were also used for characteristic wavelength extraction in subsequent oil content, linoleic acid, and lignan prediction modeling.

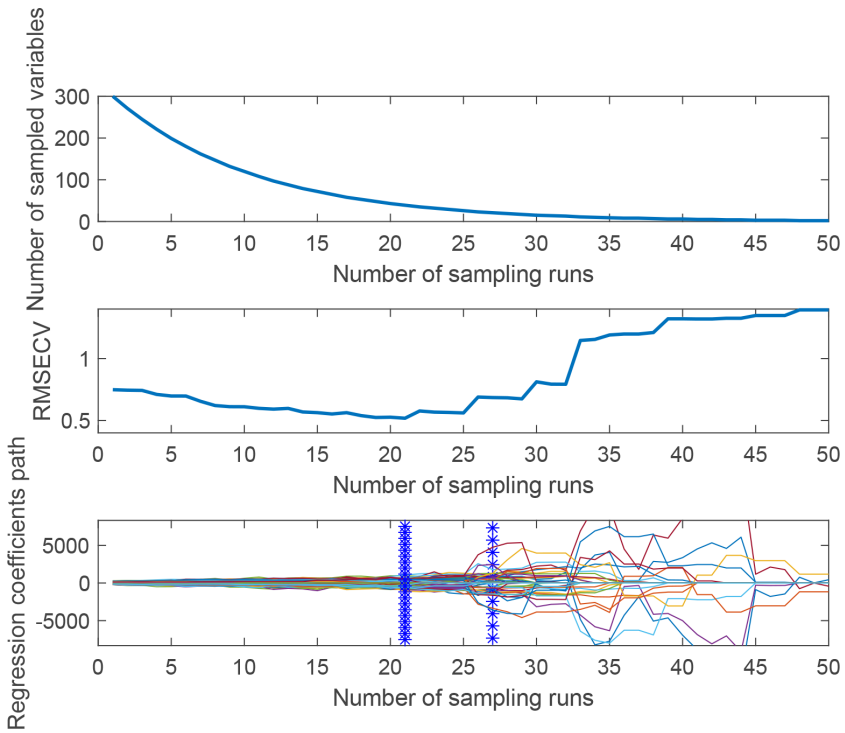


FIGURE 7
The process of extracting feature variables by CARS.

3.3 Results of modeling

3.3.1 Modeling of hyperspectral prediction of protein content in flaxseed

After determining the protein content of 30 flaxseed varieties, the original spectral data and the seven preprocessed data were combined with the actual protein content data to establish the PLSR prediction model of flaxseed protein. The cross-validation set R^2_{cv} and RMSECV were used as evaluation indexes to determine the best preprocessing method. It was found that the model prediction of the data model after SG preprocessing was the best; therefore, the SG preprocessing method was used for the original spectral data to be preprocessed. Subsequently, we utilized both feature bands and full-band data extracted from the raw bands through SPA and CARS. These data were then input into regression models, including PLSR, SVR, PCR, and MLR, to predict flaxseed protein content. The results of these predictions are presented in Table 3. An analysis of the results in Table 3 indicates that the PLSR, SVR, and PCR models, employing feature wavelengths extracted by the CARS algorithm, outperformed the models relying on full-band spectra. Specifically, they showed increased R^2_p and decreased RMSEP values. Conversely, the SPA algorithm did not enhance the predictive performance and, in some cases, even reduced it. This observation suggests that SPA trims information redundancy but may also eliminate valuable information for accurate model predictions. In summary, different algorithms extracting distinct feature wavelengths significantly influence the effectiveness of the prediction models. The optimal model, SG-CARS-MLR, exhibited a training set R^2_c of 0.9563, an RMSEC value of 0.4892%, a prediction set R^2_p of 0.9336, and an RMSEP value of 0.5616%. The results for flaxseed protein content prediction in both the training and prediction sets are illustrated in Figure 8A. The other two models, SG-CARS-PLSR and SG-CARS-PCR (Figures 8B, C), also provided reasonably accurate protein content predictions, with R^2_p values of 0.8930 and 0.8671, and RMSEP values of 0.4189% and 0.4670%, respectively. These findings confirm that the combination of HSI and the SG-CARS-MLR model delivers strong predictive performance for different flaxseed varieties' protein content. Finally, characteristic bands associated with significant protein influence were identified using the SG-CARS-MLR model (Figure 9). Generally, when the absolute t-value surpasses a specific threshold (usually 2.0), it indicates the significant impact of a corresponding independent variable on the dependent variable. In this context, Figure 8 shows that the bands at 595 and 772 nm exceed this threshold, signifying their substantial influence on the MLR model for protein content prediction.

3.3.2 Hyperspectral prediction modeling of oil content, linoleic acid and lignan in flaxseed

The prediction results for oil content, linoleic acid, and lignan content of flaxseed are presented in Table 4. The MLR model performs better than the PLSR, PCR, and SVR models. The R^2_p values of PLSR, PCR, and SVR regression algorithms are all less than 0.8, indicating these models aren't suitable for predicting the

TABLE 3 Protein prediction result table.

Modeling method	Feature extraction method	Number of feature variables	Cal		Pre	
			R^2	RMSEC	R^2	RMSEP
Protein						
PLSR	Non	320	0.9376	0.3848	0.7950	0.5800
	SPA	14	0.8933	0.5032	0.8197	0.5438
	CARS	33	0.9357	0.3907	0.8930	0.4189
SVR	Non	320	0.9546	0.3193	0.6366	0.9233
	SPA	14	0.9546	0.3193	0.6639	0.8845
	CARS	33	0.8632	0.6024	0.8061	0.7091
PCR	Non	320	0.6188	0.9512	0.4605	0.9408
	SPA	14	0.5479	1.0359	0.4282	0.9686
	CARS	33	0.9206	0.4340	0.8671	0.4670
MLR	Non	320	*	*	*	*
	SPA	14	0.9010	0.5597	0.9329	0.5642
	CARS	**33**	**0.9563**	**0.4892**	**0.9336**	**0.5616**

Represents that MLR modeling under 320 bands was not performed because MLR modeling is only applicable when the number of variables is less than the number of samples. Bold values indicate optimal model metrics.

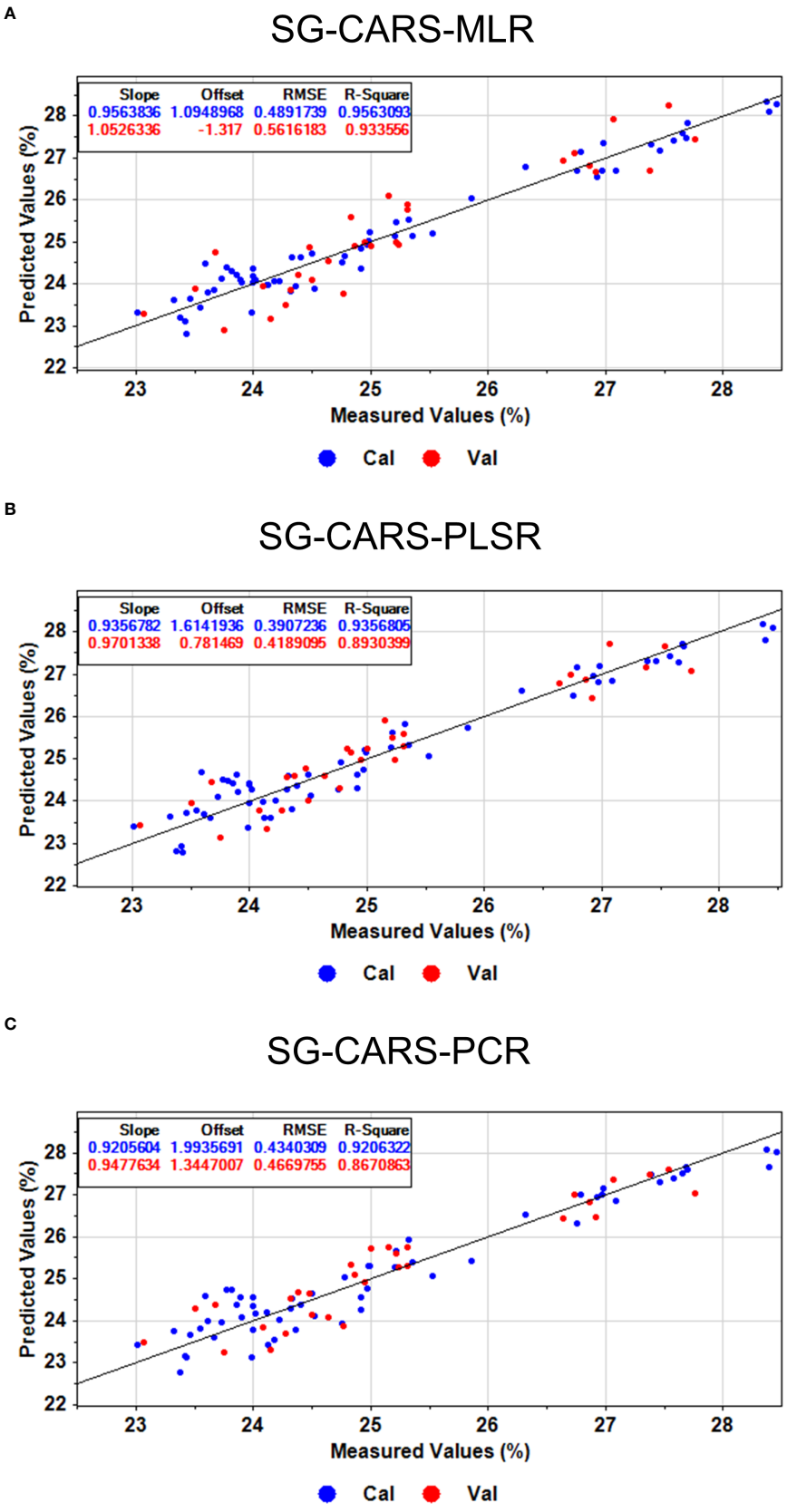


FIGURE 8
The optimal prediction of proteins based on **(A)** MLR, **(B)** PLSR, and **(C)** PCR models.

10.3389/fpls.2024.1344143

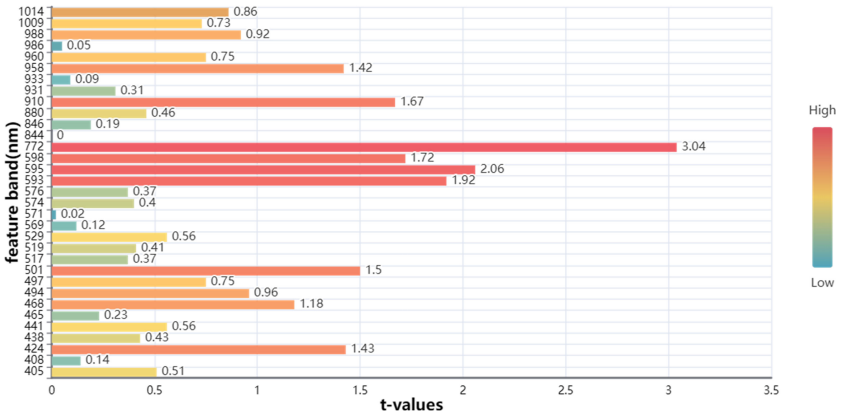


FIGURE 9
Significance map of MLR model for CARS extracted feature bands.

TABLE 4 Oil content, Linoleic acid, and lignan prediction result table.

Modeling method	Feature extraction method	Number of feature variables	Cal		Pre	
			R^2	RMSEC	R^2	RMSEP
Oil content						
PLSR	Non	320	0.7401	0.826	0.6864	0.8397
	SPA	20	0.5218	1.1205	0.6058	0.9413
	CARS	10	0.6678	0.9339	0.6438	0.8948
SVR	Non	320	0.94	0.3952	0.5884	1.0305
	SPA	20	0.9399	0.3953	0.5884	1.0305
	CARS	10	0.94	0.3953	0.5884	1.0306
PCR	Non	320	0.5835	1.0458	0.6002	0.9481
	SPA	20	0.5917	1.0353	0.6077	0.939
	CARS	10	0.6866	0.9071	0.6572	0.8779
MLR	Non	320	*	*	*	*
	SPA	**20**	**0.7675**	**0.9691**	**0.8565**	**0.8682**
	CARS	10	0.6876	1.0022	0.8532	0.8779
Linoleic acid						
PLSR	Non	320	0.7204	0.4550	0.5502	0.5497
	SPA	20	0.6871	0.4813	0.5490	0.5504
	CARS	16	0.6404	0.5160	0.4495	0.6081
SVR	Non	320	0.9461	0.1977	0.7363	0.4516
	SPA	20	0.9462	0.1977	0.7362	0.4516
	CARS	16	0.9462	0.1977	0.7362	0.4516
PCR	Non	320	0.6474	0.5110	0.5381	0.557
	SPA	20	0.4604	0.6320	0.3418	0.6649
	CARS	16	0.6564	0.5043	0.4381	0.6143
MLR	Non	320	*	*	*	*

(Continued)

TABLE 4 Continued

Modeling method	Feature extraction method	Number of feature variables	Cal		Pre	
			R^2	RMSEC	R^2	RMSEP
	SPA	**20**	**0.7489**	**0.5728**	**0.8028**	**0.5404**
	CARS	16	0.6740	0.5803	0.7286	0.6340
Lignan						
PLSR	Non	320	0.8597	0.5562	0.6626	0.8057
	SPA	29	0.5404	1.0067	0.5103	0.9707
	CARS	24	0.6362	0.8957	0.5475	0.9331
SVR	Non	320	0.9761	0.2688	0.6136	0.9082
	SPA	29	0.8464	0.6738	0.5177	1.0478
	CARS	24	0.9400	0.3953	0.5884	1.0306
PCR	Non	320	0.3959	1.1542	0.5105	0.9705
	SPA	29	0.5387	1.0086	0.4346	1.0430
	CARS	24	0.6249	0.9094	0.4880	0.9926
MLR	Non	320	*	*	*	*
	SPA	**29**	**0.9024**	**0.6562**	**0.9343**	**0.5384**
	CARS	24	0.7635	0.9455	0.8285	0.8697

Represents that MLR modeling under 320 bands was not performed because MLR modeling is only applicable when the number of variables is less than the number of samples. Bold values indicate optimal model metrics.

aforementioned contents in flaxseed. The extraction of feature wavelengths by SPA and CARS algorithms appears applicable to the MLR model. Specifically, the SG-SPA-MLR models perform better than SG-CARS-MLR in predicting oil content, linoleic acid, and lignan. For instance, the R_p^2 and RMSEP for oil content are 0.8565 and 0.8682%, and for linoleic acid are 0.8028 and 0.5404%, respectively. In contrast, the best model in literature predicting oil content for rapeseed seeds had an R_p^2 and RMSEP of 0.868 and 1.0698% (Li et al., 2023), respectively. Furthermore, lignan content was predicted with R_p^2 and RMSEP of 0.9343 and 0.5834%, respectively. Studies suggest that feature wavelengths derived from SPA and CARS algorithms enhance the predictive performance of MLR models, as observed in the prediction of moisture content of tobacco leaves (Sun et al., 2016) and the use of hyperspectral image technology for egg freshness detection (Wang et al., 2015). The scatter plots for the three types of flaxseed nutritional quality in both training and prediction sets are depicted in Figure 10, indicating the superior predictive performance of the SG-SPA-MLR model. Even though the R_p^2 for linoleic acid in the prediction set is 0.8028, the RMSEP is 0.5404%, affirming the model's aptness for prediction. Finally, Figure 11 highlights the importance of SPA-extracted feature bands in the MLR model. Figures 11A, C underscore the significance of these bands in predicting oil and lignin content. Notably, in Figure 11C, the MLR model predicts 18 feature bands with t-values greater than 2.0 in lignin content. These bands primarily appear around 470 nm (related to nitrogen content) (Li et al., 2022) and 800 nm (related to oxygen content) (Yuan et al., 2021), demonstrating the SG-SPA-MLR model's superior prediction of lignan content.

This project employs HSI technology within the 380-1018nm spectral range to gather data from flax seeds. The PLSR model cross-validation is then utilized to select the optimal pre-processing method, SG. Subsequently, characteristic wavelengths are extracted employing SPA and CARS algorithms. Finally, the spectral data corresponding to these characteristic wavelengths are combined with the protein, oil content, linoleic acid, and lignan acquired from the flax seeds through biochemical methods. This integration constructs four nutritional quality prediction models (SG-CARS/SPA-MLR) for rapid and non-destructive testing. The models achieve a prediction accuracy exceeding 0.93 for protein and lignan content, surpassing 0.85 for oil content. Although the linoleic acid content prediction accuracy is slightly lower, it still exceeds 0.80. These results fully address the requirements of practical production for rapid, non-destructive detecting of the nutritional quality of flaxseed grain.

4 Conclusions

The protein, oil content, linoleic acid, and lignan are crucial indicators for evaluating the quality of flaxseed. This study aimed to construct a model for the rapid and non-destructive detection of these components in flaxseed using HSI technology. Through experimental comparisons of various

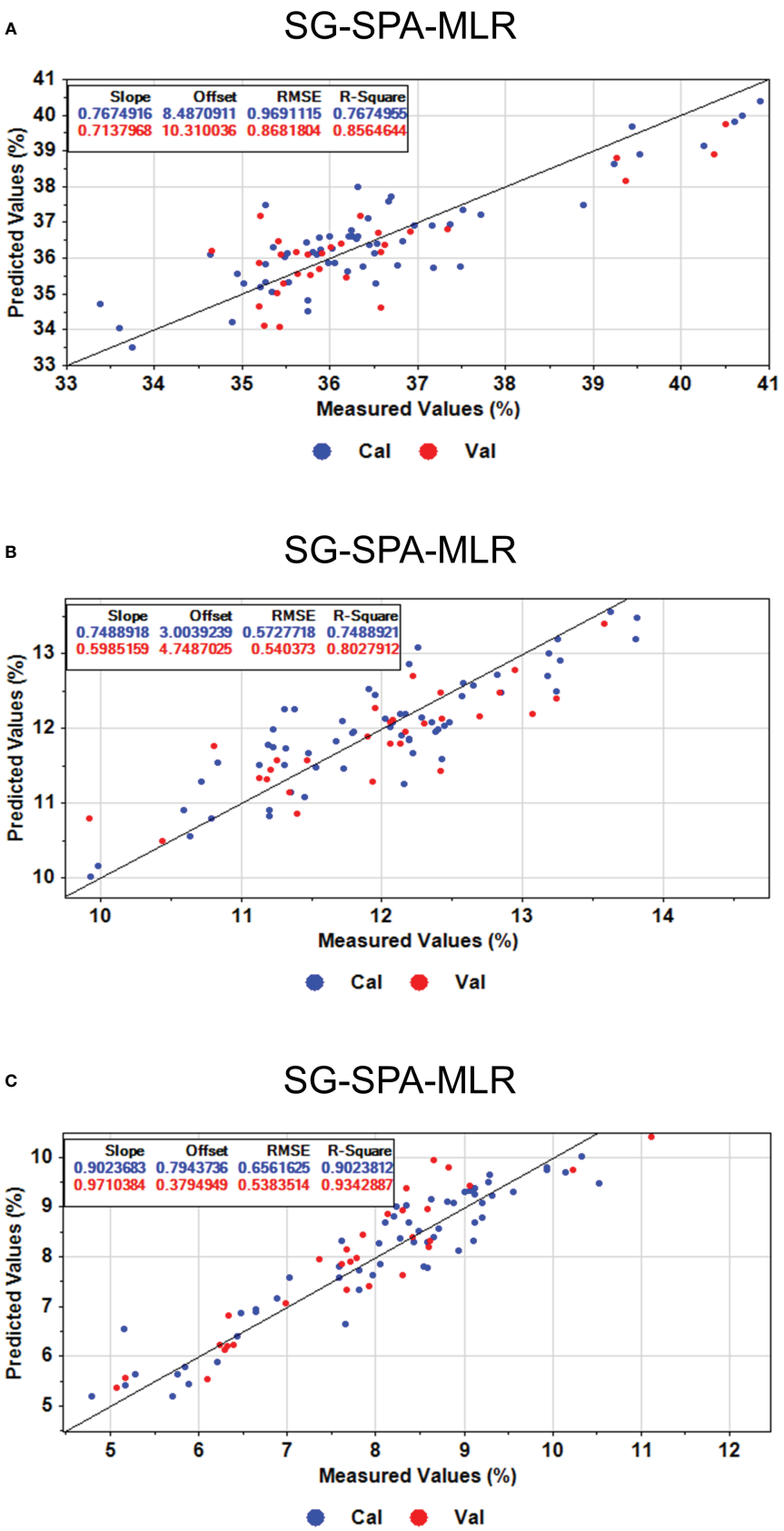


FIGURE 10
Predicted results of oil content, linoleic acid, and lignan content based on the optimal model SG-SPA-MLR. **(A)** Oil content prediction results. **(B)** Results of linoleic acid content prediction. **(C)** Prediction results of lignan content.

10.3389/fpls.2024.1344143

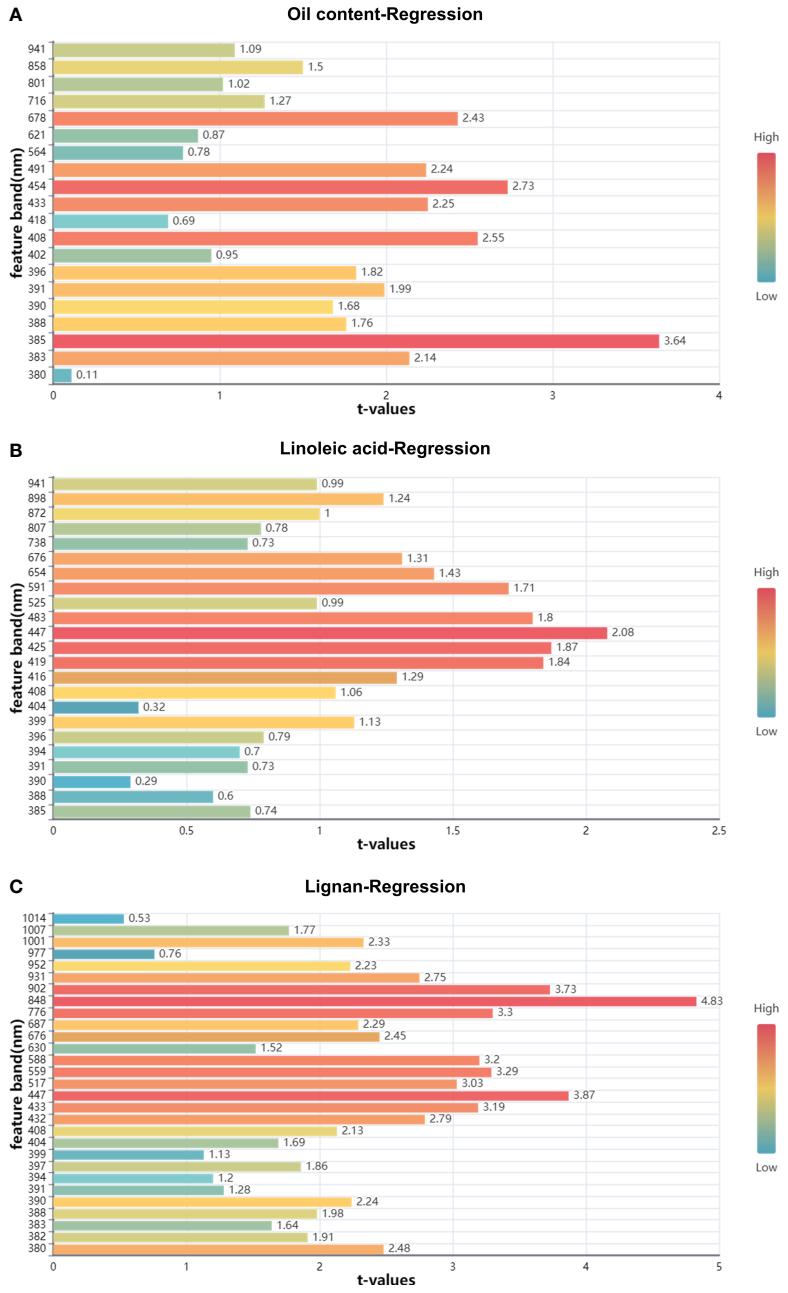


FIGURE 11
Significance map of MLR model for CARS extracted feature bands. **(A)** Significance map of the characteristic band of oil content; **(B)** Significance map of the characteristic band of linoleic acid; **(C)** Significance map of the characteristic band of lignan.

spectral image preprocessing methods and feature wavelength extraction algorithms, the preferred model achieved swift and non-destructive detection of protein, oil content, linoleic acid, and lignan in flaxseed grains, yielding better results. This research introduces a novel method for the future investigation of rapid, non-destructive, and high-precision detection of nutritional quality in different flaxseed varieties, enhancing the efficiency of screening and evaluating flax germplasm resources. The study's results hold positive practical significance for the sustainable development of the flax industry and the selection and breeding of high-quality flax varieties.

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

DZ: Conceptualization, Data curation, Methodology, Software, Writing – original draft. JH: Conceptualization, Investigation,

Resources, Supervision, Writing – review & editing. CL: Funding acquisition, Resources, Supervision, Writing – review & editing. JZ: Data curation, Resources, Writing – review & editing. YQ: Data curation, Resources, Writing – review & editing.

Funding

The author(s) declare financial support was received for the research, authorship, and/or publication of this article. National Natural Science Foundation of China (No.32360437); Innovation Fund for Higher Education of Gansu Province(No.2021A-056); Industrial Support Program for Higher Education Institutions of Gansu Province (No.2021CYZC-57).

Conflict of interest

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

Publisher's note

References

Aulia, R., Amanah, H. Z., Lee, H., Kim, M. S., Baek, I., Qin, J., et al. (2023). Protein and lipid content estimation in soybeans using Raman hyperspectral imaging. *Front. Plant Sci.* 14. doi: 10.3389/fpls.2023.1167139

Bjorgan, A., and Randeberg, L. L. (2015). *Towards real-time medical diagnostics using hyperspectral imaging technology*. Eds. J. Q. Brown and V. Deckert (Munich, Germany: European Conference on Biomedical Optics), 953712. doi: 10.1117/12.2184155

Dang, Z., and Zhao, L. (2008). Application of the near infrared reflectance spectroscopy (NIRSin analyzing flaxseed germplasm quality. *Acta Agric. Boreali-Occident. Sin.* 17, 110–113.

Feng, X., Li, G., Song, J., and Shao, H. (2016). Determination of lignan in flaxseed by HPLC. *Anim. Husb. Feed Sci.* 37, 17–18+22. doi: 10.16003/j.cnki.issn1672-5190.2016.12.005

Goyal, A., Sharma, V., Upadhyay, N., Gill, S., and Sihag, M. (2014). Flax and flaxseed oil: an ancient medicine & modern functional food. *J. Food Sci. Technol.* 51, 1633–1653. doi: 10.1007/s13197-013-1247-9

Hu, H., Wang, T., Wei, Y., Xu, Z., Cao, S., Fu, L., et al. (2023). Non-destructive prediction of isoflavone and starch by hyperspectral imaging and deep learning in Puerariae Thomsonii Radix. *Front. Plant Sci.* 14. doi: 10.3389/fpls.2023.1271320

Jiang, M., Li, Y., Song, J., Wang, Z., Zhang, L., Song, L., et al. (2023). Study on black spot disease detection and pathogenic process visualization on winter jujubes using hyperspectral imaging system. *Foods* 12, 435. doi: 10.3390/foods12030435

Katare, C., Saxena, S., Agrawal, S., Prasad, G., and Bisen, P. S. (2012). Flax seed: a potential medicinal food. *J. Nutr. Food Sci.* 2, 120–127.

Li, H., Liang, Y., Xu, Q., and Cao, D. (2009). Key wavelengths screening using competitive adaptive reweighted sampling method for multivariate calibration. *Anal. Chim. Acta* 648, 77–84. doi: 10.1016/j.aca.2009.06.046

Li, X., Peng, F., Wei, Z., Han, G., and Liu, J. (2023). Non-destructive detection of protein content in mulberry leaves by using hyperspectral imaging. *Front. Plant Sci.* 14. doi: 10.3389/fpls.2023.1275004

Li, M., Zhu, X., Li, W., Tang, X., Yu, X., and Jiang, Y. (2022). Retrieval of nitrogen content in apple canopy based on unmanned aerial vehicle hyperspectral images using a modified correlation coefficient method. *Sustainability* 14, 1992. doi: 10.3390/su14041992

Liu, F., Jin, Z. L., Naeem, M. S., Tian, T., Zhang, F., He, Y., et al. (2011). Applying near-infrared spectroscopy and chemometrics to determine total amino acids in herbicide-stressed oilseed rape leaves. *Food Bioprocess Technol.* 4, 1314–1321. doi: 10.1007/s11947-010-0445-y

Lu, Y., Jia, B., Yoon, S.-C., Zhuang, H., Ni, X., Guo, B., et al. (2022). Spatio-temporal patterns of Aspergillus flavus infection and aflatoxin B1 biosynthesis on maize kernels probed by SWIR hyperspectral imaging and synchrotron FTIR micro spectroscopy. *Food Chem.* 382, 132340. doi: 10.1016/j.foodchem.2022.132340

Ma, C., Ren, Z., Zhang, Z., Du, J., Jin, C., and Yin, X. (2021). Development of simplified models for nondestructive testing of rice (with husk) protein content using hyperspectral imaging technology. *Vib. Spectrosc.* 114, 103230. doi: 10.1016/j.vibspec.2021.103230

Ma, J., Sun, D.-W., Pu, H., Cheng, J.-H., and Wei, Q. (2019). Advanced techniques for hyperspectral imaging in the food industry: principles and recent applications. *Annu. Rev. Food Sci. Technol.* 10, 197–220. doi: 10.1146/annurev-food-032818-121155

Mahesh, S., Jayas, D. S., Paliwal, J., and White, N. D. G. (2015). Comparison of partial least squares regression (PLSR) and principal components regression (PCR) methods for protein and hardness predictions using the near-infrared (NIR) hyperspectral images of bulk samples of Canadian wheat. *Food Bioprocess Technol.* 8, 31–40. doi: 10.1007/s11947-014-1381-z

Meng, D., Ji, Z., and Ren, L. (2001). Determination of linoleic acid in grain by underivatized high performance liquid chromatography. *Anal. Test. Technol. Instrum.* 71–74.

Mueller, K., Eisner, P., and Kirchhoff, E. (2010). Simplified fractionation process for linseed meal by alkaline extraction – Functional properties of protein and fiber fractions. *J. Food Eng.* 99, 49–54. doi: 10.1016/j.jfoodeng.2010.01.036

Oomah, B. D. (2001). Flaxseed as a functional food source. *J. Sci. Food Agric.* 81, 889–894. doi: 10.1002/jsfa.898

Ozaki, Y. (2021). Infrared spectroscopy—Mid-infrared, near-infrared, and far-infrared/terahertz spectroscopy. *Anal. Sci.* 37, 1193–1212. doi: 10.2116/analsci.20R008

Rajkumar, P., Wang, N., EImasry, G., Raghavan, G. S. V., and Gariepy, Y. (2012). Studies on banana fruit quality and maturity stages using hyperspectral imaging. *J. Food Eng.* 108, 194–200. doi: 10.1016/j.jfoodeng.2011.05.002

Ribeiro, L. F., Peralta-Zamora, P. G., Maia, B. H. L. N. S., Ramos, L. P., and Pereira-Netto, A. B. (2013). Prediction of linolenic and linoleic fatty acids content in flax seeds and flax seeds flours through the use of infrared reflectance spectroscopy and multivariate calibration. *Food Res. Int.* 51, 848–854. doi: 10.1016/j.foodres.2013.01.061

Shao, Y., Wang, Y., Xuan, G., Gao, C., Wang, K., and Gao, Z. (2020). Visual detection of SSC and firmness and maturity prediction for feicheng peach by using hyperspectral imaging. *Trans. Chin. Soc Agric. Mach.* 51, 344–350.

Sun, J., Zhou, X., Wu, X., Zhang, X., and Li, Q. (2016). Identification of moisture content in tobacco plant leaves using outlier sample eliminating algorithms and hyperspectral data. *Biochem. Biophys. Res. Commun.* 471, 226–232. doi: 10.1016/j.bbrc.2016.01.125

Tian, X., Fan, S., Huang, W., Wang, Z., and Li, J. (2020). Detection of early decay on citrus using hyperspectral transmittance imaging technology coupled with principal component analysis and improved watershed segmentation algorithms. *Postharvest Biol. Technol.* 161, 111071. doi: 10.1016/j.postharvbio.2019.111071

Tu, K., Wen, S., Cheng, Y., Xu, Y., Pan, T., Hou, H., et al. (2022). A model for genuineness detection in genetically and phenotypically similar maize variety seeds based on hyperspectral imaging and machine learning. *Plant Methods* 18, 81. doi: 10.1186/s13007-022-00918-7

Wang, Y., Guo, W., Zhu, X., and Liu, Q. (2019). Effect of homogenization on detection of milk protein content based on NIR diffuse reflectance spectroscopy. *Int. J. Food Sci. Technol.* 54, 387–395. doi: 10.1111/ijfs.13948

Wang, Z., Huang, W., Tian, X., Long, Y., Li, L., and Fan, S. (2022). Rapid and non-destructive classification of new and aged maize seeds using hyperspectral image and chemometric methods. *Front. Plant Sci.* 13. doi: 10.3389/fpls.2022.849495

Wang, Q., Zhou, K., Wang, C., and Ma, M. (2015). Egg freshness detection based on hyperspectral image technology. *Adv. J. Food Sci. Technol.* 7, 652–657. doi: 10.19026/ajfst.7.1623

Xiang, Y., Chen, Q., Su, Z., Zhang, L., Chen, Z., Zhou, G., et al. (2022). Deep learning and hyperspectral images based tomato soluble solids content and firmness estimation. *Front. Plant Sci.* 13. doi: 10.3389/fpls.2022.860656

Xu, Y., Wu, W., Chen, Y., Zhang, T., Tu, K., Hao, Y., et al. (2022). Hyperspectral imaging with machine learning for non-destructive classification of Astragalus membranaceus var. mongholicus, Astragalus membranaceus, and similar seeds. *Front. Plant Sci.* 13. doi: 10.3389/fpls.2022.1031849

Yadav, P. K., Burks, T., Frederick, Q., Qin, J., Kim, M., and Ritenour, M. A. (2022). Citrus disease detection using convolution neural network generated features and Softmax classifier on hyperspectral image data. *Front. Plant Sci.* 13. doi: 10.3389/fpls.2022.1043712

Yang, J., Sun, L., Xing, W., Feng, G., Bai, H., and Wang, J. (2021). Hyperspectral prediction of sugarbeet seed germination based on gauss kernel SVM. *Spectrochim. Acta A. Mol. Biomol. Spectrosc.* 253, 119585. doi: 10.1016/j.saa.2021.119585

Yao, S., Liao, M., Kang, J., Wei, Z., Liu, N., and REn, H. (2022). Optimization of simultaneous extraction of oil, protein and gum from flaxseed by enzyme – assisted three phase partitioning. *China Oils Fats* 47, 11–17. doi: 10.19902/j.cnki.zgyz.1003-7969.210239

Ye, J., Jia, H., Guo, D., Yan, W., and Xie, L. (2021). Establishment and applicant of near-infrared reflectance spectroscopy models for predicting protein, linolenic acid and lignan contents of flaxseed. *Chin. J. Oil Crop Sci.* 43, 353–360. doi: 10.19802/j.issn.1007-9084.2019308

Yoosefzadeh-Najafabadi, M., Earl, H. J., Tulpan, D., Sulik, J., and Eskandari, M. (2021). Application of machine learning algorithms in plant breeding: predicting yield from hyperspectral reflectance in soybean. *Front. Plant Sci.* 11. doi: 10.3389/fpls.2020.624273

Yu, H., Liu, H., Wang, N., Yang, Y., Shi, A., Liu, L., et al. (2016). Rapid and visual measurement of fat content in peanuts by using the hyperspectral imaging technique with chemometrics. *Anal. Methods* 8, 7482–7492. doi: 10.1039/C6AY02029A

Yu, K.-Q., Zhao, Y.-R., Liu, Z.-Y., Li, X.-L., Liu, F., and He, Y. (2014). Application of visible and near-infrared hyperspectral imaging for detection of defective features in loquat. *Food Bioprocess Technol.* 7, 3077–3087. doi: 10.1007/s11947-014-1357-z

Yuan, Z., Ye, Y., Wei, L., Yang, X., and Huang, C. (2021). Study on the optimization of hyperspectral characteristic bands combined with monitoring and visualization of pepper leaf SPAD value. *Sensors* 22, 183. doi: 10.3390/s22010183

Zhang, Y., and Guo, W. (2020). Moisture content detection of maize seed based on visible/near-infrared and near-infrared hyperspectral imaging technology. *Int. J. Food Sci. Technol.* 55, 631–640. doi: 10.1111/ijfs.14317

Zhang, H., Hou, Q., Luo, B., Tu, K., Zhao, C., and Sun, Q. (2022). Detection of seed purity of hybrid wheat using reflectance and transmittance hyperspectral imaging technology. *Front. Plant Sci.* 13. doi: 10.3389/fpls.2022.1015891

Zhang, Z.-S., Wang, L.-J., Li, D., Li, S.-J., and Özkan, N. (2011). Characteristics of flaxseed oil from two different flax plants. *Int. J. Food Prop.* 14, 1286–1296. doi: 10.1080/10942911003650296

Zhu, S., Chao, M., Zhang, J., Xu, X., Song, P., Zhang, J., et al. (2019). Identification of soybean seed varieties based on hyperspectral imaging technology. *Sensors* 19, 5225. doi: 10.3390/s19235225

Zou, Z., Chen, J., Wu, W., Luo, J., Long, T., Wu, Q., et al. (2023). Detection of peanut seed vigor based on hyperspectral imaging and chemometrics. *Front. Plant Sci.* 14. doi: 10.3389/fpls.2023.1127108

frontiers | Frontiers in Plant Science

TYPE Original Research
PUBLISHED 24 May 2024
DOI 10.3389/fpls.2024.1351301

OPEN ACCESS

EDITED BY
Xi Tian,
Beijing Academy of Agriculture and Forestry Sciences, China

REVIEWED BY
Magdi A. A. Mousa,
King Abdulaziz University, Saudi Arabia
Xia Liang,
The University of Melbourne, Australia
Xiaoyu Tian,
Jiangsu University, China

*CORRESPONDENCE
Pei Wang
peiwang@swu.edu.cn

RECEIVED 06 December 2023
ACCEPTED 10 May 2024
PUBLISHED 24 May 2024

CITATION
Li H, Wan L, Li C, Wang L, Zhu S, Chen X and Wang P (2024) Hyperspectal imaging technology for phenotyping iron and boron deficiency in *Brassica napus* under greenhouse conditions.
Front. Plant Sci. 15:1351301.
doi: 10.3389/fpls.2024.1351301

Hyperspectal imaging technology for phenotyping iron and boron deficiency in *Brassica napus* under greenhouse conditions

Hui Li[1,2], Long Wan[1], Chengsong Li[1,3], Lihong Wang[1], Shiping Zhu[1], Xinping Chen[4] and Pei Wang[1,2,5]*

[1]College of Engineering and Technology, Key Laboratory of Agricultural Equipment for Hilly and Mountain Areas, Southwest University, Chongqing, China, [2]Key Laboratory of Modern Agricultural Equipment and Technology (Jiangsu University), Ministry of Education, School of Agricultural Engineering, Jiangsu University, Zhenjiang, China, [3]National Citrus Engineering Research Center, Chinese Academy of Agricultural Sciences & Southwest University, Chongqing, China, [4]Interdisciplinary Research Center for Agriculture Green Development in Yangtze River Basin, College of Resources and Environment, Southwest University, Chongqing, China, [5]Institute of Urban Agriculture, Chinese Academy of Agricultural Sciences, Chengdu, China

Introduction: The micronutrient deficiency of iron and boron is a common issue affecting the growth of rapeseed (*Brassica napus*). In this study, a non-destructive diagnosis method for iron and boron deficiency in *Brassica napus* (genotype: Zhongshuang 11) using hyperspectral imaging technology was established.

Methods: The recognition accuracy was compared using the Fisher Linear Discriminant Analysis (LDA) and Support Vector Machine (SVM) recognition models. Recognition results showed that Multiple Scattering Correction (MSC) could be applied for the full band hyperspectral data processing, while the LDA models presented better performance on establishing the leaf iron and boron deficiency symptom recognition than the SVM models.

Results: The recognition accuracy of the training set reached 96.67%, and the recognition rate of the prediction set could be 91.67%. To improve the model accuracy, the Competitive Adaptive Reweighted Sampling algorithm (CARS) was added to construct the MSC-CARS-LDA model. 33 featured wavelengths were selected via CARS. The recognition accuracy of the MSC-CARS-LDA training set was 100%, while the recognition accuracy of the MSC-CARS-LDA prediction set was 95.00%.

Discussion: This study indicates that, it is capable to identify the iron and boron deficiency in rapeseed using hyperspectral imaging technology.

KEYWORDS

deficiency identification, rapeseed, iron and boron, hyperspectral imaging, MSC-CARS-LDA

1 Introduction

Rapeseed, rich in oil and protein, is an important oil crop and industrial raw material, as well as a potential bioenergy crop overall the world (Liu et al., 2019). According to the Rapeseed Explorer of USDA, the global production of rapeseed has reached 87.103 million metric tons, which producing 31.8 million metric rapeseed oil (USDA, 2024). With various useful compounds of fatty acids, vitamins and proteins, rapeseed oil ranks as the third most popular vegetable oil after oil palm and soybean (Friedt et al., 2018). To better manage the fertilizer supply during rapeseed cultivation, it is essential to monitoring the micronutrients status of the plants.

Appropriate application content of fertilizers will not only benefit the absorption and utilization of nutrients by the crop plants, but also contribute to plants stress tolerance (Hasanuzzaman et al., 2018; Thor, 2019). On the contrary, lacking essential nutrients could inhibit the growth of the plants, which would directly lead to the negative effect on rapeseed quality or yield (Agren et al., 2012; Johnson et al., 2022). Real-time, fast, and accurate monitoring of nutrient content would provide guidance for reasonable fertilization to increase the crop quality and production in any specific regions (Brown et al., 2022; Tian et al., 2024). Therefore, monitoring the nutrient content of plants is an important aspect in crop cultivation and management.

As two of the essential micronutrients, boron and iron play important roles in the growth and reproduction of rapeseed, especially in the southwestern region of China. Boron participates in promoting the transport of carbohydrates *in vivo* plants which will accelerate the growth of apical shoots and meristem. It is also conducive to the development of plant flower organs (Kalaji et al., 2018; Li et al., 2020). When the rapeseed plant is in deficiency of boron, the transportation of assimilation products *in vivo* plant could be interrupted. As a result, a large amount of starch would accumulate in the leaves and petioles. Furtherly, the greatly increase of phenolic compounds content would lead to necrosis of plant apical buds. Thus, the main manifestation of boron deficiency in rapeseed is the inhibition of apical buds, which would interrupt the growth of roots and shoots, ultimately leading to the issue of "blooming but not setting fruit" of rapeseed plants. Iron acts as an activator of some enzymes or enzyme cofactors in the synthesis of chlorophyll. It would indirectly affect the production process of chlorophyll, while playing an important role in electron transfer chain in various biochemical reactions *in vivo* plants (Takano et al., 2008; Pavlovic et al., 2021). The main manifestation of iron deficiency in rapeseed is the chlorosis and yellowing between leaf veins while the leaf veins themselves remain green, especially in the top fresh leaves (Takano et al., 2008; Merchant, 2010).

Phenotyping technology with optical sensors such as RGB camera, chlorophyll fluorescence sensors, and particularly the spectral imaging system, has been widely applied in monitoring various biotic stresses for crops. With UAV-based RGB and multi-spectral sensors, salinity stress phenotyping has been realized in tomato and quinoa plants (Johansen et al., 2019; Jiang et al., 2022). The study results provided insight into the effects of salt stress on plant area, growth and condition. Optical information like chlorophyll fluorescence can also be measured for photosystem status evaluation such as investigating herbicide stresses in soybean plants (Li et al., 2018). Meanwhile, phenotyping of stresses from over or deficient macronutrients such as nitrogen, phosphorus, and potassium have also been successfully tested in many studies using hyperspectral imaging technology (Jiang et al., 2015; Tmušić et al., 2020). However, at present, the deficiency of the micronutrients iron and boron in crops is mainly evaluated using artificial vision and empirical morphological diagnostic methods which could only be made when obvious stress symptoms have appeared, and the specific fertilization may be missed for the suitable application time window.

The objectives of this study were to, (1) investigate if it was possible to differentiate the iron and boron deficiency symptoms in rapeseed from healthy plants at early growth stage using spectral imaging technology; (2) optimize the spectral diagnostic model for a high classification accuracy. The results will provide support for the nutritional diagnosis of iron and boron content in rapeseed fields using UAV-based sensing systems and even for the potential detection of vegetation deficiency symptoms via the space-based remote sensing satellites.

2 Materials and methods

2.1 Plant materials

The cultivar of tested rapeseed (*Brassica napus* L.) in this study is Zhongshuang 11 (ZS11, Beijing, Chinese Academy of Agricultural Sciences), which is widely grown in the Yangtze River basin. The plants were grown in 380 mm×300 mm pots with soilless hydroponic incubator with six plants per pot. All the plants were cultivated in a greenhouse of Southwest University in Chongqing, China.

The nutrient deficient plants were cultivated as the methodology described by Han et al. (2016), in which the ZS11 genotype was also cultivated as the tested plants. The rapeseed seeds were germinated in distilled water. After germination, the seedlings were transferred to a plastic net floating on the half strength modified Hoagland solution (Table 1). Normally growth seedlings were selected for next cultivation steps of the tests. Seedlings for control treatment were kept in the half strength modified Hoagland solution with boron concentration of 20 μmol L^{-1} (H_3BO_3) and iron concentration of 80 μmol L^{-1} ($C_{10}H_{12}FeN_2NaO_8$, EDTA-Fe), which were dramatically lower than the element concentration in the Chinese State Standard of foliar microelement fertilizer (State Administration for Marker Regulation of the People's Republic of China & Standardization Administration of the People's Republic of China, 2020). The seedlings for nutrient deficiency treatments were then transferred to solution with iron or boron in lower concentration. Boron deficiency plants was transferred to the solution with boron concentration of 0.5 μmol L^{-1}, while the iron deficiency plants were treated with iron concentration of 1 μmol L^{-1}. The other nutrients of both micronutrients deficient solution were kept in same

10.3389/fpls.2024.1351301

TABLE 1 Modified 1/2 Hoagland complete nutrient solution formula.

Chemical	Molecular Weight	Concentration (10^{-3} mol L^{-1})
$Ca(NO_3)_2{\cdot}4H_2O$	236.15	2500
KNO_3	101.1	2500
NH_4NO_3	80.04	1000
K_2SO_4	174.26	250
$MgSO_4{\cdot}7H_2O$	246.47	1000
KH_2PO_4	136.09	500
DETA-Fe	376.05	80
H_3BO_3	61.83	20
$MnCI_2{\cdot}4H_2O$	197.91	4.5
$ZnSO_4{\cdot}7H_2O$	287.54	0.3
$CuSO_4{\cdot}5H_2O$	249.68	0.16
$(NH4)_6Mo_7O_{24}{\cdot}4H_2O$	1235.86	0.16

concentration as the half strength modified Hoagland solution. The solution in all treatments was replaced every two days.

Germination treatment was applied to full and consistent ZS11 seeds. The seeds were soaked in distilled water for 20 minutes and disinfected with 5% NaClO solution for 20 minutes. After rinsed with distilled water repeatedly for 5–6 times, the seeds were put on gauze soaked in 1/4 strength Hoagland solution for seedling cultivation.

Seedlings with uniform growth stage were selected for transplant. The plants were transferred to plastic hydroponic tanks containing nutrient solutions (1/4 strength Hoagland solution was used for cultivation in the first week after transplanting, half strength Hoagland solution was used for cultivation since the second week after transplanting, solution for nutrient deficiency treatments were applied since the third week after transplanting). Four biological replicates were applied for each treatment with 72 plants in total. The plants were set with a Randomized Block Design. The experiment was repeated twice in 3rd March to 14th June and 7th September to 12th December in 2023.

2.2 Hyperspectral imaging system

The physical and architectural diagrams of the hyperspectral imaging system are shown in Figures 1A, B, respectively. The main hardware includes a hyperspectral camera (Raptor EM285CL, Raptor Photonics Led., UK), a spectrometer (Impector V10E, Measuring wavelength range 364~1025 nm, Spectral resolution 2.8 nm, Specim, Spectral Imaging Ltd., Finland), a zoom lens, a 150 W halogen adjustable light source, a linear photoconductor, a stepping motor mobile platform, a computer, etc. The whole set of devices is placed in the black box except the computer. The main software installed on the computer includes Spectral image, an image acquisition software provided by Wuling Optics (Taiwan, China), and HIS Analyzer, an image analysis software.

This device was fixed on the top of a black box in a darkroom. Each plant was moved to the measuring platform out from the hydroponic incubator. The window of the black box was closed during measurement.

2.3 Data collection and calibration

After 28 days of transplanting, hyperspectral images of rapeseed leaves were collected uniformly. To ensure the representativeness of the collected data, all samples were placed horizontally under the same conditions for imaging. After pre-testing, it was ultimately determined that the exposure time of the hyperspectral imaging system camera was 48 ms, the working distance from the lens to the sample was 480 mm, and the moving platform speed was 1.12 mm s^{-1}.

Black and white board correction was performed on the hyperspectral image data of each sample in the image analysis software HIS Analyzer. The correction formula is as follows:

$$R = (R_s - R_D)/(R_w - R_D)$$

where, R is the relative reflection density of the leaves, R_s is the reflection density of the original image of the sample, R_w is the reflection density of all the white calibration image, and R_D is the reflection density of the all black calibration image. Black and white correction is used to eliminate the influence of camera dark current, while converting the spectral values of the original hyperspectral image into reflectance.

2.4 Data preprocessing

Due to the influence of instruments, image acquisition background, environmental lighting conditions, and other factors, there would be noise, spectral baseline drift, and translation in the obtained spectral data. To eliminate these adverse effects on classification modeling, preprocessing of spectral data is necessary. After preliminary experiments, normalization, SG convolutional smoothing, spectral differentiation, and Multiple Scattering Correction (MSC) were selected for the spectra preprocess of the leaf samples after smoothing. The preprocessing procedure is shown in Figure 2. Four types of the spectral data obtained after preprocessing are spectral sample sets 1–4, which are abbreviated as RAW, 1st Der, MSC, and 2nd Der in the following text.

Figure 3 shows four spectral samples obtained from the pre-processed spectral data of some healthy rapeseed leaves.

2.5 Band screening

Hyperspectral data often has hundreds or even thousands of wavelength points, which not only provides rich information about samples but also poses challenges for computer storage, transmission, and data processing (Arnon and Hoagland, 1938). When extracting spectral dimension information from hyperspectral data for modeling, using full band spectral

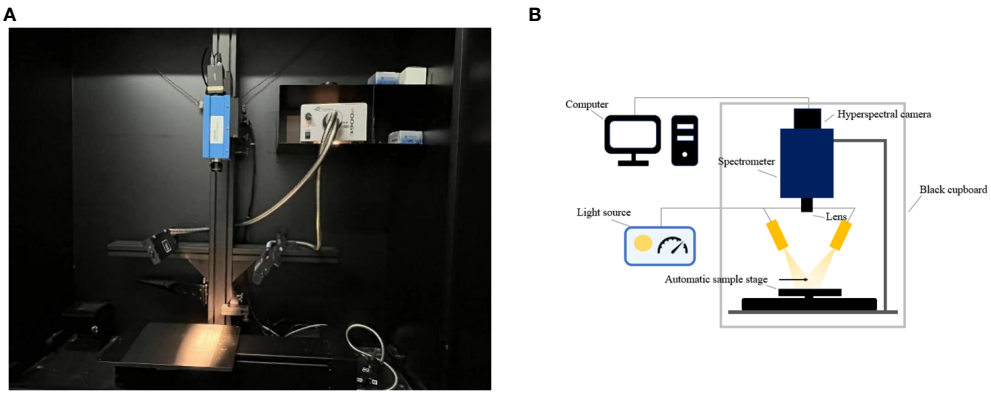


FIGURE 1
Hyperspectral imaging system **(A)** physical drawing; **(B)** architecture diagram.

information to establish the model will bring various negative impacts to the model due to the presence of uninformative variables in the data (Gruber et al., 2013). The dimension reduction algorithm can select the wavelength variables that are more meaningful to the classification results from the full wavelength range and eliminate redundant wavelengths. It could improve the prediction accuracy and modeling calculation efficiency of the model, as well as reducing the overfitting of the model and improve the generalization ability of the model (Arnon and Hoagland, 1938; Gruber et al., 2013; Khan et al., 2018).

2.5.1 Continuous projection algorithm

Successive Projections Algorithm (SPA) is a forward variable selection algorithm, which uses vector projection analysis to select the combination of many variables with the smallest collinearity. In some studies on plant spectral feature classification and regression models, continuous projection algorithms are often applied in the dimensionality reduction process of hyperspectral data, which can play a good role in improving model operation efficiency and recognition accuracy (Belgiu and Drăguţ, 2016).

2.5.2 Competitive adaptive reweighting algorithm

The Competitive Adaptive Reweighted Sampling (CARS) algorithm has also been widely applied in the recognition of plant spectral features. CARS uses the Monte Carlo sampling principle to select sample subsets for modeling, and to evaluate the importance of variables through the regression coefficients of the sub models. In each iteration, dimensionality reduction is achieved by removing variables with smaller mean regression coefficients through Exponential Decreasing Function (EDF) and Adaptive Reweighted Sampling (ARS) (Lorente et al., 2012).

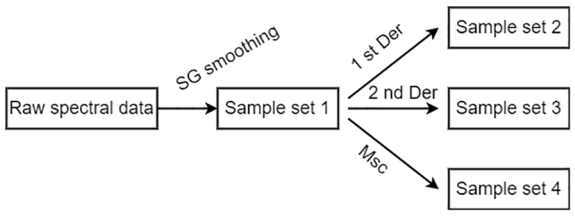


FIGURE 2
Flow chart of spectral data preprocessing.

2.6 Classification model

Linear Discriminant Analysis (LDA), also known as Fisher linear discriminant analysis, is a classic algorithm for pattern recognition and is widely used in multi class classification problems. Using LDA can maximize the inter class scatter matrix of the projected pattern samples and minimize the intra class scatter matrix, ensuring that the projected pattern samples have the minimum intra class distance and maximum inter class distance in a new space. Its essence lies in finding a subspace. It enables better separation of various categories in this subspace, which means that patterns have the best separability in that space (Zhang et al., 2022).

Support Vector Machine (SVM) is a supervised pattern recognition method. The original spectral data is mapped to a high-dimensional feature space, and an optimal classification hyperplane is constructed to maximize the distance between the support vectors of various samples and this hyperplane. SVM can be used for linear and nonlinear multivariate analysis problems, and the support vector can be solved by using linear equations instead of Quadratic programming. By selecting appropriate kernel functions to ensure the speed and efficiency of modeling while implementing nonlinear mapping (Yuan et al., 2020), this experiment uses Radial Basis Function (RBF) as kernel functions.

3 Results

3.1 Spectral features of nutrient deficient leaves

From the average spectra of the collected leaves of rapeseed plants (Figure 4), it presented that overall waveform of the spectral

10.3389/fpls.2024.1351301

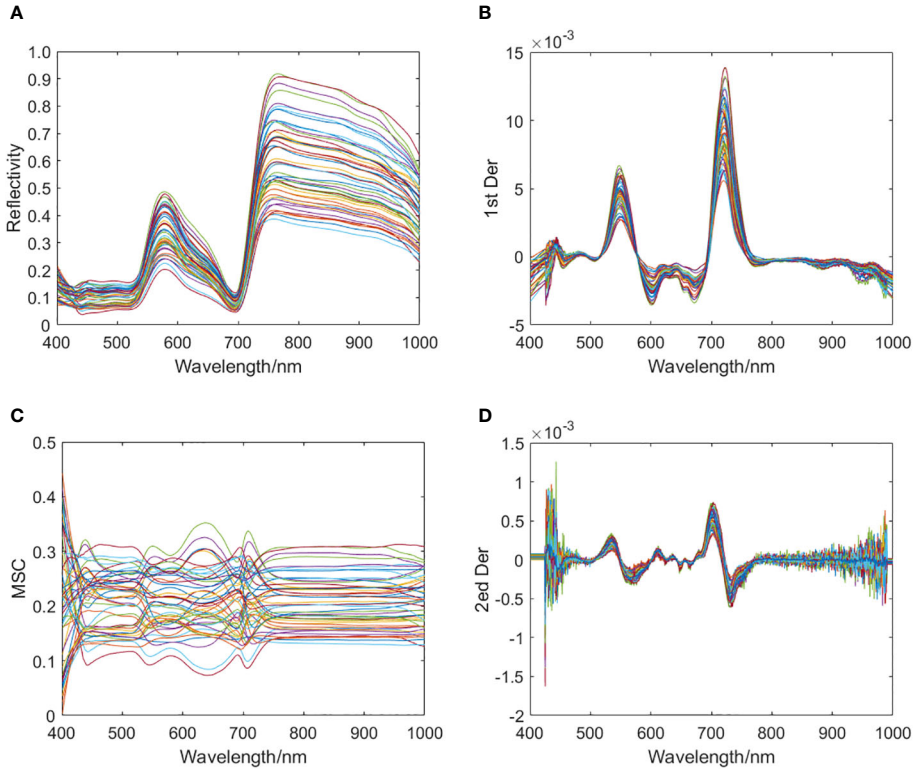


FIGURE 3
Four spectral samples obtained from preprocessing spectral data of some healthy rapeseed leaves. **(A)** Sample set 1: RAW; **(B)** Sample set 2: 1st Der; **(C)** Sample set 3: MSC; **(D)** Sample set 4: 2nd Der.

reflection curve in the wavelength range of 400–1000 nm was similar between normal plants and iron or boron deficient plants.

3.2 Deficiency recognition and classification model based on full band information

After preprocessing the 400–1000 nm full band spectral data using three preprocessing methods, known as spectral first order differential, spectral second order differential and MSC. LDA discriminative model and SVM discriminative model for identifying iron deficiency, boron deficiency and normal leaves were established respectively. LDA discriminative model is a typical Fisher linear discriminant analysis in Matlab Toolbox. When using SVM to build a discriminative model, the kernel function used when using SVM to build a discriminative model is the radial basis function (RBF) kernel function:

$$K(X_i, X_j) = \exp(-\gamma \parallel X_i - X_j \parallel)^2$$

In the SVM modeling, the Penalty coefficient y was set as 100, and the kernel width σ was set as 0.1. The discrimination results of each model were shown in Table 2.

Comparing the discrimination accuracy of the two models, it presented that the LDA model had better overall discrimination performance than the SVM model. However, when using the LDA model to model the rapeseed leaf spectral dataset, the most suitable preprocessing method was MSC. When using the SVM model, the two preprocessing methods MSC and 2nd Der had better results.

By analyzing the confusion matrix of the modeling set (Figure 5) and the test set (Figure 6) based on the SVM model, it presented that the SVM model had a good spectral recognition effect for healthy and nutrient deficient rape leaves, with an accuracy rate of more than 90%. The recognition effect of iron

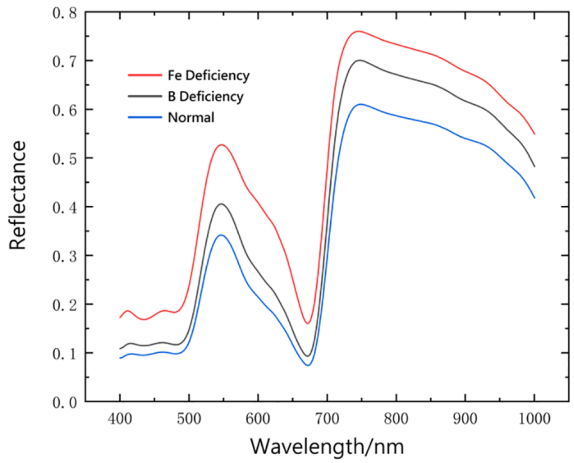


FIGURE 4
Comparison of average spectra between normal leaves and iron and boron deficient in *Brassica napus* leaves.

10.3389/fpls.2024.1351301

TABLE 2 Discrimination results of LDA and SVM discrimination models under different spectral pretreatment.

Model	Preprocessing method	Training set		Testing set	
		Nc/Nt	Accuracy	Nc/Nt	Accuracy
SVM	RAW	81/90	90.00%	47/60	78.33%
	1st Der	78/90	86.67%	39/60	65.00%
	MSC	84/90	93.33%	52/60	86.67%
	2nd Der	83/90	92.22%	51/60	85.00%
LDA	RAW	79/90	87.77%	49/60	81.66%
	1st Der	81/90	90.00%	46/60	76.67%
	MSC	87/90	96.67%	55/60	91.67%
	2nd Der	84/90	93.33%	47/60	78.33%

"Nc" represents the correct discriminant number of the tested samples; "Nt" represents the total number of tested samples.

and boron deficient rapeseed leaves is average, with the 1st Der data having the worst effect, with an accuracy rate of only 55%. The accuracy rates of RAW, MSC, and 2nd Der data are all between 70% and 85%.

Experiments also shown that SVM models based on full band spectral data can effectively identify healthy rapeseed leaves and rapeseed leaves lacking iron and boron elements. However, the recognition accuracy between iron deficient and boron deficient leaves still needed to be improved.

The analysis of the confusion matrix of the modeling set (Figure 7) and the test set (Figure 8) based on the LDA model showed that the LDA model was superior to the SVM model in spectral recognition of healthy and nutrient deficient rapeseed leave. Its accuracy in the test set is more than 95%. In the recognition of iron and boron deficient rapeseed leaves, MSC data showed significantly better performance than RAW, 1st Der, and 2nd Der data, with average accuracy exceeding 90% in both training and testing sets.

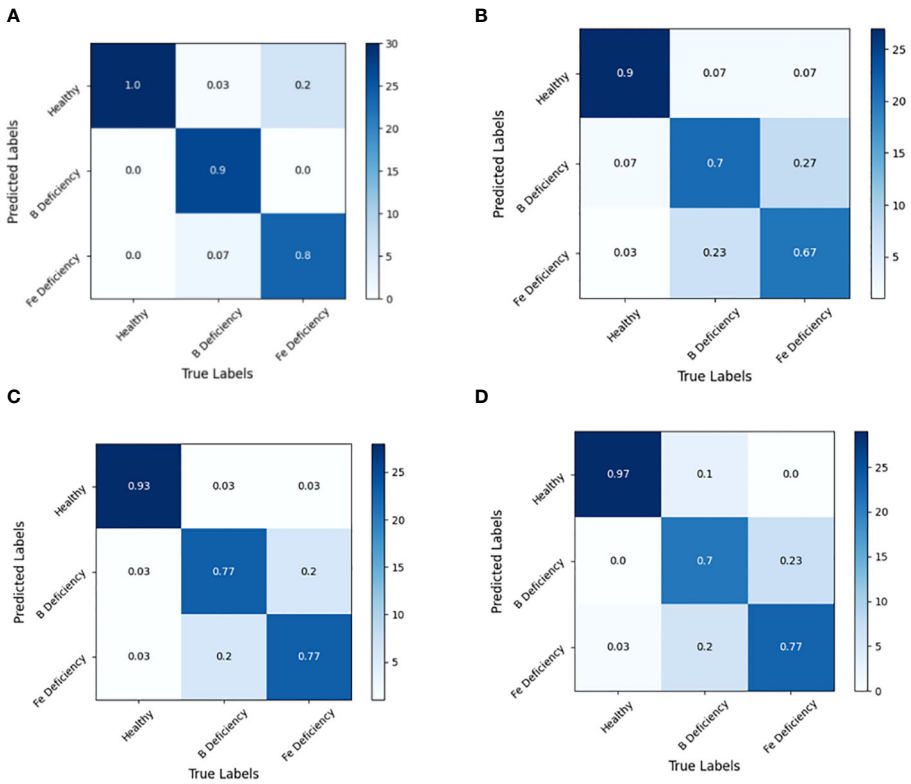


FIGURE 5
Discrimination results based on SVM models under different spectral pretreatments (Modeling set) for iron and boron deficiency in *Brassica napus*. **(A)** RAW, **(B)** 1st Der, **(C)** MSC, **(D)** 2nd Der.

10.3389/fpls.2024.1351301

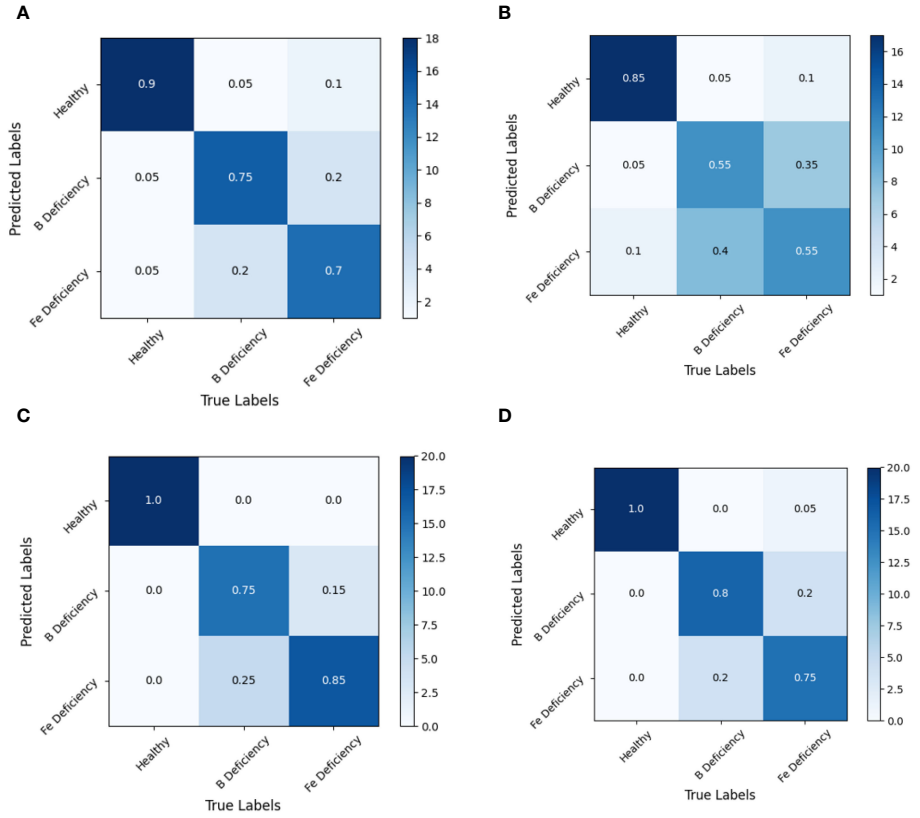


FIGURE 6
Discrimination results based on SVM model under different spectral pretreatments (Test set) for iron and boron deficiency in *Brassica napus*. **(A)** RAW, **(B)** 1st Der, **(C)** MSC, **(D)** 2nd Der.

Further analysis of the discrimination results of the MSC-LDA model showed that both the modeling and prediction sets had a 100% accuracy in discriminating normal samples. In general, the discrimination accuracy of samples with boron deficiency was higher than that of samples with iron deficiency. From the confusion matrix, two samples with boron deficiency in the modeling set were wrongly identified as samples with iron deficiency symptoms, while one sample with iron deficiency symptoms was wrongly identified as samples with boron deficiency, and two samples are wrongly identified as healthy samples. The prediction set discrimination results also showed that one sample with boron deficiency was wrongly identified as samples with iron deficiency symptoms, and four samples with iron deficiency symptoms were wrongly identified as samples with iron deficiency.

The results of this experiment indicate that the MSC-LDA model achieved the highest accuracy in the combination of data preprocessing and modeling methods for Brassica napus iron and boron. The overall discrimination accuracy of the modeling set reached 96.67%, and the overall discrimination accuracy of the prediction set reached 91.67%.

3.3 Feature band screening results

Figure 9A shows the process of reducing the number of bands involved in modeling through 50 Monte Carlo sampling (MC) of the sample data. Figure 9B shows the cross-validation error curve of the PLS model using the Leave on One Out (LOO) method as the number of bands involved in modeling decreases. From the above two curves, it could be seen that as the number of bands involved in modeling gradually decreases, the Root Mean Square Error of Cross Validation (RMSECV) of the model first shows a slow decreasing trend. It indicated that there is indeed a lot of redundant information in the spectral raw data containing more than 600 bands. Screening out certain band data could not only reduce computational complexity, but also improve the accuracy of the model to a certain extent. When the sampling frequency starts from 24, the RMSECV of the model in the training set gradually increased as the number of modeling bands decreases, indicating that some band data useful for classification modeling begins to be eliminated.

The above phenomenon indicates that there is indeed a large amount of redundant information in the original spectrum that is useless for the classification and recognition of iron and boron stress in rapeseed. It is meaningful to reduce the dimensionality of the original spectral data.

SPA and CARS were used to reduce the dimensionality of rapeseed leaf spectral data, as shown in Table 3. A total of 18 characteristic wavelengths were selected by SPA and defined as subset 1. CARS screened a total of 33 wavelengths and defined them as subset 2.

10.3389/fpls.2024.1351301

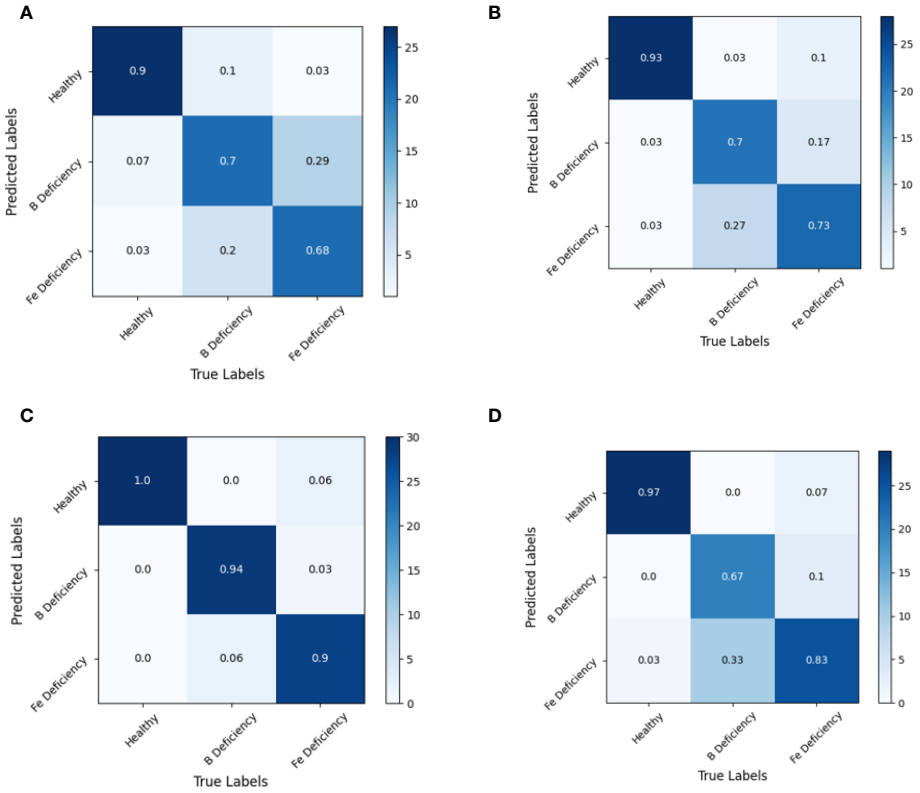


FIGURE 7
Discrimination results based on LDA models under different spectral pretreatments (Modeling set) for iron and boron deficiency in *Brassica napus*. **(A)** RAW, **(B)** 1st Der, **(C)** MSC, **(D)** 2nd Der.

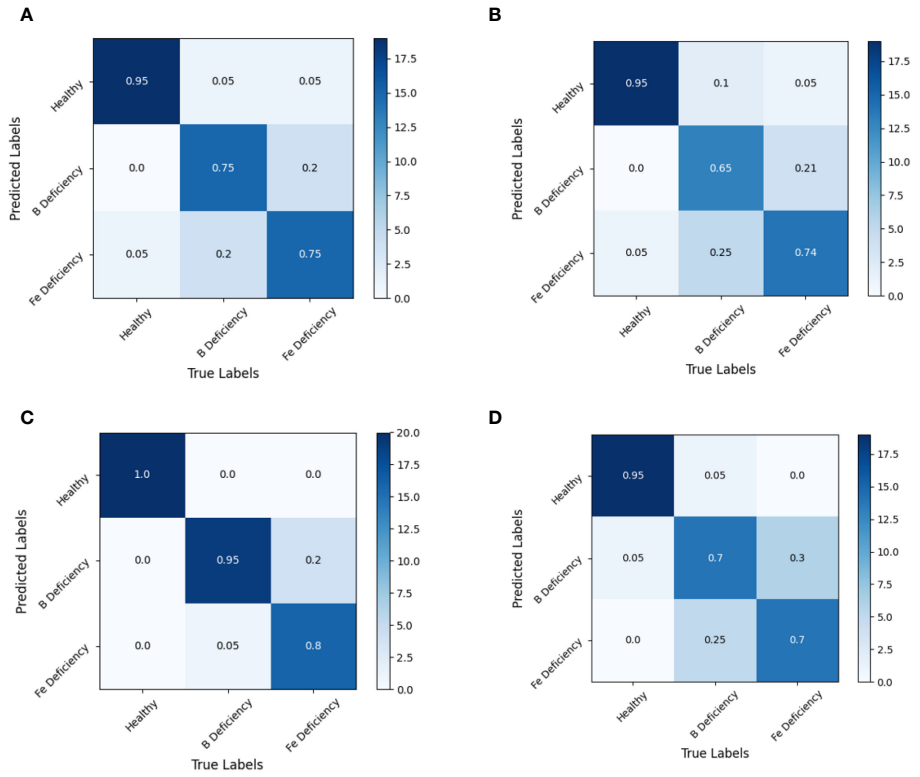


FIGURE 8
Discrimination results based on LDA model under different spectral pretreatment (test set). **(A)** RAW, **(B)** 1st Der, **(C)** MSC, **(D)** 2nd Der.

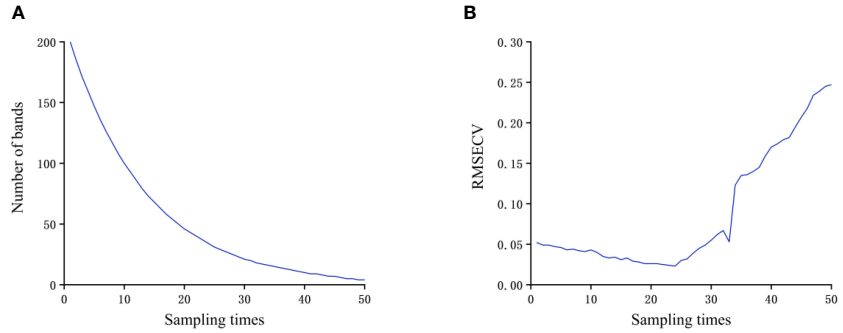


FIGURE 9
Impact of the number of bands involved in modeling on model accuracy. **(A)** Change trend of band number; **(B)** Model RMSECV change trend.

3.4 Establishment of a deficiency recognition and classification model based on feature band information

Since the MSC-LDA model is superior to other discriminative model when full band spectral information modeling is used, the MSC-SPA-LDA and MSC-CARS-LDA modeling and discrimination are conducted using two characteristic wavelength subsets screened according to SPA and CARS. The prediction results are shown in Table 4, and the confusion matrix is shown in Figure 10 (MSC-SPA-LDA) and Figure 11 (MSC-CARS-LDA).

Analysis of the discrimination results of the MSC-SPA-LDA model showed that the accuracy of the MSC-SPA-LDA model based on SPA feature bands is 94.44% on the training set and 91.67% on the test set, which is slightly lower than the MSC-LDA model based on full band. However, due to its significant reduction in the number of input variables in the model, the running speed of the model is significantly improved, and the accuracy is within an acceptable range, so the MSC-SPA-LDA model has better applicability than the MSC-LDA model.

By analyzing the discrimination results of the MSC-CARS-LDA model, it could be concluded that the MSC-CARS-LDA model based on CARS feature bands achieved 100% and 95% accuracy on the training and testing sets, respectively, making it the model with the highest recognition accuracy in this experiment.

The feature wavelengths of subset 2 selected based on the CARS algorithm were mainly concentrated between the regions of 400–450 nm and 800–1000 nm, especially in the blue-violet light region of 400–450 nm, which is not the green peak region with the greatest difference in the 500–650 nm spectral curve. This indicates that the degree of leaf chlorosis is not the only basis for discrimination in this recognition system.

TABLE 3 Selected characteristic wavebands by SPA and CARS.

Data name	Wavelength(nm)
Dataset 1	410, 411, 416, 421, 426, 429, 434, 437, 441, 442, 672, 691, 722, 737, 981, 990, 997
Dataset 2	401, 402, 405, 406, 408, 411, 412, 414, 417, 418, 422, 426, 427, 429, 430, 434, 438, 445, 446, 449, 453, 455, 674, 688, 812, 864, 882, 919, 955, 973, 977, 980, 992

4 Discussion

The deficiency of iron and boron could lead to a decrease in chlorophyll content in the leaves, weakening their absorption of solar radiation, and causing an overall increase in the spectral reflectance of the leaves in the wavelength range of 400–700 nm, resulting in a "blue shift" phenomenon at the "red edge" position. This was consistent with previous research results (Yang et al., 2018). Considering the "green peak" at 550 nm, the spectral reflectance difference was the largest. The increase in green peak caused by iron deficiency was more intense than that caused by boron deficiency, indicating that the level of plant nutrient element content was closely related to spectral characteristics. When the plants were in deficiency of iron or boron, the total chlorophyll content of their leaves might reduce. That would lead to weak absorption of solar radiation and an increase in the reflectance and transmittance of incident light, which has been proven in crops like sorghum and sugar beet (Teixeira et al., 2020; Wu et al., 2021). The symptoms of nutrient deficiency in rapeseed leaves appeared because of the decrease of chlorophyll content, which might cause corresponding spectral responses such as an increase in green peaks. This provides a basis for conducting spectral recognition and identification.

The spectral reflectance of plant leaves in the range of 400–1000 nm indicated spectral responses to various factors such as plant metabolites, chlorophyll, water content, internal structure of leaf surfaces, and physical properties of plant leaves. The correlation between spectral reflectance of different bands and the abundance or deficiency of iron and boron elements in rapeseed plants was comprehensive responses of the rapeseed plants to the nutrients status and environment, rather than the direct correlation between spectral values and iron and boron content. Therefore, machine learning algorithms was employed in this study for further analysis. Two pattern recognition methods, LDA and SVM, were used to

10.3389/fpls.2024.1351301

TABLE 4 MSC-LDA discrimination model based on characteristic wavelength.

Model	Training Set		Test Set	
	Nc/Nt	Accuracy	Nc/Nt	Accuracy
MSC-SPA-LDA	85/90	94.44%	55/60	91.67%
MSC-CARS-LDA	90/90	100.00%	57/60	95.00%

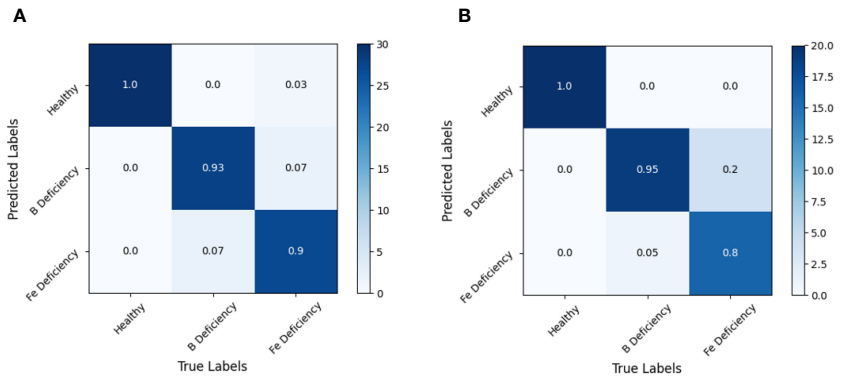


FIGURE 10
MSC-SPA-LDA model discrimination result confusion matrix. **(A)** training set, **(B)** test set.

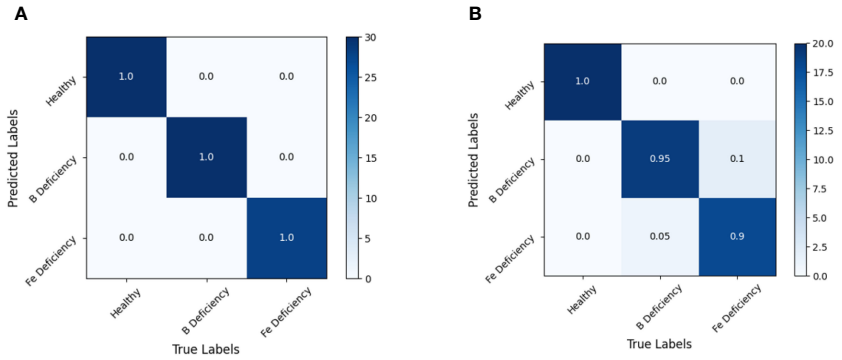


FIGURE 11
MSC-CARS-LDA model discrimination result confusion matrix. **(A)** training set, **(B)** test set.

identify different deficiency symptoms. The LDA algorithm achieves better recognition results, and the CARS algorithm performs better than the SPA algorithm in feature wavelength screening. Through the analysis of confusion matrix, it presented that the recognition rate of the recognition model established in this study for healthy rape leaves was always higher than 90%. The recognition of iron deficient leaves and boron deficient leaves presented some confusion of samples. From Figure 5, it could also be seen that the green peak increase response caused by iron deficiency was stronger than that caused by boron deficiency. It might suggest that the physiological response to spectral properties from iron deficiency was more sensitive than that from boron deficiency in rapeseed plants (Sarafi et al., 2018).

5 Conclusion

The spectral response of normal, iron deficient, and boron deficient rapeseed plants was investigated using hyperspectral imaging technology in this study. Thus, it could conclude that, (1) with employing spectral imaging technology, it is capable to identify the iron and boron deficiency symptoms in rapeseed from healthy plants at early growth stage based on full band and featured band; (2) the LDA discriminative model established by screening characteristic wavelengths could be optimized using CARS for further field application with lower data consumption and faster calculation, and the recognition accuracy of its modeling set and prediction set could be 92.22% and 86.67%.

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

HL: Data curation, Funding acquisition, Investigation, Software, Writing – original draft, Writing – review & editing. LW: Data curation, Investigation, Writing – original draft, Writing – review & editing. CL: Resources, Writing – review & editing. LHW: Resources, Supervision, Writing – review & editing. SZ: Methodology, Resources, Supervision, Validation, Writing – review & editing. XC: Methodology, Resources, Supervision, Writing – review & editing. PW: Data curation, Funding acquisition, Investigation, Methodology, Software, Validation, Writing – original draft, Writing – review & editing.

Funding

The author(s) declare financial support was received for the research, authorship, and/or publication of this article. This research was funded by the National Natural Science Foundation of China, grant number 32201651 and 32001425; the Fundamental Research Funds for the Central Universities (SWU-KT22024); the Natural Science Foundation of Chongqing, China, grant numbers cstc2020jcyj-msxmX0414 and cstc2020jcyj-msxmX0459; the Open Funding of the Key Laboratory of Modern Agricultural Equipment and Technology (Jiangsu University), grant numbers MAET202105; Local Financial of National Agricultural Science & Technology Center, Chengdu, grant number NASC2021KR02.

Acknowledgments

The authors would like to appreciate Prof. Dr. Shouyong Xie, Prof. Dr. Yuheng Yang, Dr. Shengling Zhou, Dr. Xiaozhong Wang, Dr. Dunyi Liu, Dr. Qi Niu, Dr. Chunlei Wang, Dr. Wei Ma and Mr. Linyun Chen for technical support.

Conflict of interest

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

Publisher's note

References

Agren, G. I., Wetterstedt, J. A. M., and Billberger, M. F. K. (2012). Nutrient limitation on terrestrial plant growth–modeling the interaction between nitrogen and phosphorus. *New Phytol.* 194, 953–960. doi: 10.1111/j.1469-8137.2012.04116.x

Arnon, D. I., and Hoagland, D. R. (1938). Water-culture method for growing plants without soil, California Agricultural Experiment Station. (California, USA: California Agricultural Experiment Station).

Belgiu, M., and Drăguţ, L. (2016). Random forest in remote sensing: A review of applications and future directions. *ISPRS J. photogrammetry Remote Sens.* 114, 24–31. doi: 10.1016/j.isprsjprs.2016.01.011

Brown, P. H., Zhao, F. J., and Dobermann, A. (2022). What is a plant nutrient? Changing definitions to advance science and innovation in plant nutrition. *Plant Soil* 476, 11–23. doi: 10.1007/s11104-021-05171-w

Friedt, W., Tu, J., and Fu, T. (2018). Academic and economic importance of *Brassica napus* rapeseed. In: S. Liu, R. Snowdon and B. Chalhoub. *Brassica napus Genome*. Cham: Springer, 1–20. doi: 10.1007/978-3-319-43694-4_1

Gruber, B. D., Giehl, R. F. H., Friedel, S., and von Wirén, N. (2013). Plasticity of the Arabidopsis root system under nutrient deficiencies. *Plant Physiol.* 163, 161–179. doi: 10.1104/pp.113.218453

Han, P., Qin, L., Li, Y., Liao, X., Xu, Z., Yu, C., et al. (2016). Effects of different nutrient deficiencies on growth and root morphological changes of rapeseed seedlings (*Brassica napus* L.). *Chin. J. Oil Crop Sci.* 38, 88. doi: 10.7505/j.issn.1007-9084.2016.01.014

Hasanuzzaman, M., Bhuyan, M. B., Nahar, K., Hossain, M. S., Mahmud, J. A., Hossen, M. S., et al. (2018). Potassium: A vital regulator of plant responses and tolerance to abiotic stresses. *Agronomy* 8, 31. doi: 10.3390/agronomy8030031

Jiang, J., Johansen, K., Stanschewski, C. S., Wellman, G., Mousa, M. A., Fiene, G. M., et al. (2022). Phenotyping a diversity panel of quinoa using UAV-retrieved leaf area index, SPAD-based chlorophyll and a random forest approach. *Precis. Agric.* 23, 961–983. doi: 10.1007/s11119-021-09870-3

Jiang, L., He, X., Wei, J., Wang, X., and Wang, F. (2015). Advances of phytochemicals in mediating micronutrients metabolisms. *Chin. Bull. Life Sci.* 27, 1020–1027. doi: 10.13376/j.cbls/2015141

Johansen, K., Morton, M. J., Malbeteau, Y. M., Aragon, B., Al-Mashharawi, S. K., Ziliani, M. G., et al. (2019). Unmanned aerial vehicle-based phenotyping using morphometric and spectral analysis can quantify responses of wild tomato plants to salinity stress. *Front. Plant Sci.* 10, 370. doi: 10.3389/fpls.2019.00370

Johnson, R., Vishwakarma, K., Hossen, M. S., Kumar, V., Shackira, A. M., Puthur, J. T., et al. (2022). Potassium in plants: Growth regulation, signaling, and environmental stress tolerance. *Plant Physiol. Biochem.* 172, 56–69. doi: 10.1016/j.plaphy.2022.01.001

Kalaji, H. M., Bąba, W., Gediga, K., Goltsev, V., Samborska, I. A., Cetner, M. D., et al. (2018). Chlorophyll fluorescence as a tool for nutrient status identification in rapeseed plants. *Photosynthesis Res.* 136, 329–343. doi: 10.1007/s11120-017-0467-7

Khan, M. J., Khan, H. S., Yousaf, A., Khurshid, K., and Abbas, A. (2018). Modern trends in hyperspectral image analysis: A review. *IEEE Access* 6, 14118–14129. doi: 10.1109/ACCESS.2018.2812999

Li, L., Chen, B., Yan, G., Gao, G., Xu, K., Xie, T., et al. (2020). Proposed strategies and current progress of research and utilization of oilseed rape germplasm in China. *J. Plant Genet. Resour.* 21, 1–19. doi: 10.13430/j.cnki.jpgr.20200109005

Li, H., Wang, P., Weber, J. F., and Gerhards, R. (2018). Early identification of herbicide stress on soybean (*Glycine max* (L.) Merr.) using chlorophyll fluorescence imaging technology in field practice. *Sensors* 18, 21. doi: 10.3390/s18010021

Liu, C., Feng, Z., Xiao, T., Ma, X., Zhou, G., Huang, F., et al. (2019). Development, potential and adaptation of Chinese rapeseed industry. *Chin. J. Oil Crop Sci.* 41, 485. doi: 10.7505/j.issn.1007-9084.2019.04.001

Lorente, D., Aleixos, N., Gómez-Sanchis, J., Cubero, S., García-Navarrete, O. L., and Blasco, J. (2012). Recent advances and applications of hyperspectral imaging for fruit and vegetable quality assessment. *Food Bioprocess Technol.* 5, 1121–1142. doi: 10.1007/s11947-011-0725-1

Merchant, S. S. (2010). The elements of plant micronutrients. *Plant Physiol.* 154, 512–515. doi: 10.1104/pp.110.161810

Pavlovic, J., Kostic, L., Bosnic, P., Kirkby, E. A., and Nikolic, M. (2021). Interactions of silicon with essential and beneficial elements in plants. *Front. Plant Sci.* 12, 697592. doi: 10.3389/fpls.2021.697592

Sarafi, E., Siomos, A., Tsouvaltzis, P., Therios, I., and Chatzissavvidis, C. (2018). Boron toxicity effects on the concentration of pigments, carbohydrates and nutrient elements in six non-grafted pepper cultivars (*Capsicum annuum* L.). *Indian J. Plant Physiol.* 23, 474–485. doi: 10.1007/s40502-018-0388-2

State Administration for Marker Regulation of the People's Republic of China & Standardization Administration of the People's Republic of China (2020). *Foliar microelement fertilizer: GB/T 17420—2020* Vol. 2020 (Beijing: Standards Press of China).

Takano, J., Miwa, K., and Fujiwara, T. (2008). Boron transport mechanisms: collaboration of channels and transporters. *Trends Plant Sci.* 13, 451–457. doi: 10.1016/j.tplants.2008.05.007

Teixeira, G. C. M., de Mello Prado, R., Oliveira, K. S., D'Amico-Damião, V., and da Silveira Sousa Junior, G. (2020). Silicon increases leaf chlorophyll content and iron nutritional efficiency and reduces iron deficiency in sorghum plants. *J. Soil Sci. Plant Nutr.* 20, 1311–1320. doi: 10.1007/s42729-020-00214-0

Thor, K. (2019). Calcium-nutrient and messenger. *Front. Plant Sci.* 10, 440. doi: 10.3389/fpls.2019.00440

Tian, X., Yao, J., Yu, H., Wang, W., and Huang, W. (2024). Early contamination warning of Aflatoxin B1 in stored maize based on the dynamic change of catalase activity and data fusion of hyperspectral images. *Comput. Electron. Agric.* 217, 108615.

Tmušić, G., Manfreda, S., Aasen, H., James, M. R., Gonçalves, G., Ben-Dor, E., et al. (2020). Current practices in UAS-based environmental monitoring. *Remote Sens.* 12, 1001. doi: 10.3390/rs12061001

USDA (2024) Rapeseed Explorer. Available online at: https://ipad.fas.usda.gov/cropexplorer/cropview/commodityView.aspx?cropid=2226000.

Wu, Z., Wang, X., Song, B., Zhao, X., Du, J., and Huang, W. (2021). Responses of photosynthetic performance of sugar beet varieties to foliar boron spraying. *Sugar Tech* 23, 1332–1339. doi: 10.1007/s12355-021-01008-z

Yang, Y., Wu, Y., Li, W., Liu, X., Zheng, J., Zhang, W., et al. (2018). Determination of geographical origin and icariin content of *Herba Epimedii* using near infrared spectroscopy and chemometrics. *Spectrochimica Acta Part A: Mol. Biomolecular Spectrosc.* 191, 233–240. doi: 10.1016/j.saa.2017.10.019

Yuan, Z., Wei, L., Zhang, Y., Yu, M., and Yan, X. (2020). Hyperspectral inversion and analysis of heavy metal arsenic content in farmland soil based on optimizing cars combined with PSO-SVM algorithm. *Spectrosc. Spectral Anal.* 40, 567–573. doi: 10.3964/j.issn.1000-0593(2020)02-0567-07

Zhang, N., Zhang, X., Wang, C., Li, L., and Bai, T. (2022). Cotton LAI estimation based on hyperspectral and successive projection algorithm. *Trans. Chin. Soc. Agric. Machinery* 53, 257–262. doi: 10.6041/j.issn.1000-1298.2022.S1.028

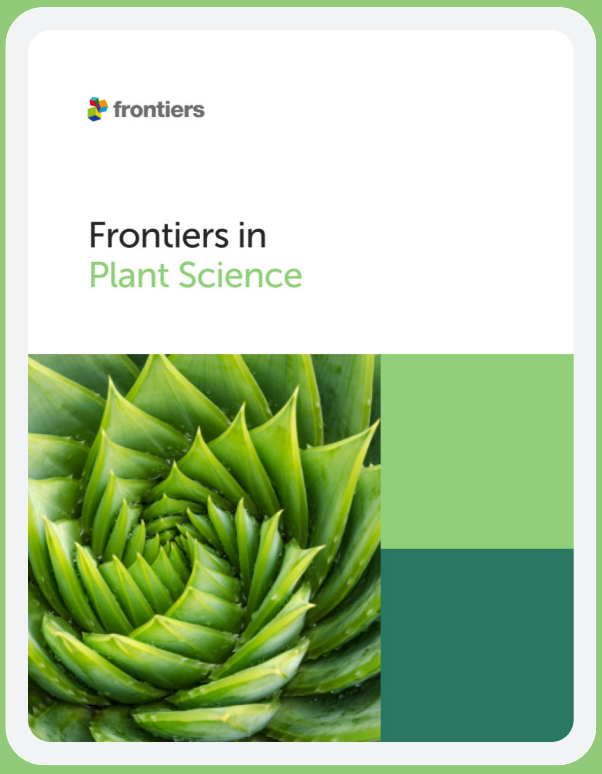
Frontiers in
Plant Science

Cultivates the science of plant biology and its applications

The most cited plant science journal, which advances our understanding of plant biology for sustainable food security, functional ecosystems and human health.

Discover the latest Research Topics

See more →



Frontiers

Avenue du Tribunal-Fédéral 34
1005 Lausanne, Switzerland
frontiersin.org

Contact us

+41 (0)21 510 17 00
frontiersin.org/about/contact