



OPEN ACCESS

EDITED BY

Qingli Dong,
University of Shanghai for Science and
Technology, China

REVIEWED BY

LiFei Huang,
Guangdong Academy of Agricultural
Sciences, China
Qiang Guan,
Inner Mongolia University for Nationalities,
China

*CORRESPONDENCE

Hongseok Lee
✉ ehg117@korea.kr

RECEIVED 01 October 2025

REVISED 30 November 2025

ACCEPTED 01 December 2025

PUBLISHED 15 December 2025

CITATION

Kang T-A, An C-Y, Jeong H, Park Y-J, Park J,
Ryu J, Bae H, Yoon Y and Lee H (2025) Early
detection of sesame leaf disease using
hyperspectral imaging and machine learning:
a patch-based spatial-spectral integration
approach.
Front. Sustain. Food Syst. 9:1716907.
doi: 10.3389/fsufs.2025.1716907

COPYRIGHT

© 2025 Kang, An, Jeong, Park, Park, Ryu, Bae,
Yoon and Lee. This is an open-access article
distributed under the terms of the [Creative
Commons Attribution License \(CC BY\)](#). The
use, distribution or reproduction in other
forums is permitted, provided the original
author(s) and the copyright owner(s) are
credited and that the original publication in
this journal is cited, in accordance with
accepted academic practice. No use,
distribution or reproduction is permitted
which does not comply with these terms.

Early detection of sesame leaf disease using hyperspectral imaging and machine learning: a patch-based spatial-spectral integration approach

Tae-An Kang^{1,2}, Chi-Yong An¹, Hoejeong Jeong¹, You-Jin Park¹,
Jinki Park¹, Jongsoo Ryu¹, Hyungjin Bae¹, Youngnam Yoon¹ and
Hongseok Lee^{1*}

¹Department of Upland Crop Research and Development, National Institute of Crop Science, Rural Development Administration, Miryang, Gyeongsangnam-do, Republic of Korea, ²Department of Applied Biosciences, Kyungpook National University, Daegu, Republic of Korea

Introduction: Sesame (*Sesamum indicum* L.) is highly susceptible to leaf blight caused by *Corynespora cassicola* and bacterial leaf spot caused by *Pseudomonas syringae*, both of which can cause substantial yield and quality losses. Conventional detection methods are limited to post-symptom diagnosis, limiting opportunities for timely intervention. Early asymptomatic detection is essential for stable production and economic sustainability.

Methods: We developed a hyperspectral imaging framework integrated with machine learning algorithms for early sesame leaf diseases detection by implementing a 3 × 3 patch-based approach to capture spatial-spectral disease patterns. Hyperspectral images (450–950 nm) were acquired using dedicated hyperspectral imaging equipment. The 3 × 3 patch extraction method captured spatial-spectral disease progression patterns, including infected areas and surrounding healthy tissues, to enable effective early disease detection. Three preprocessing methods [Raw reflectance spectra, Savitzky–Golay (SG) smoothing, and 1st derivative] were evaluated. Wavelength selection was performed using the Sequential Projection Algorithm (SPA), Genetic Algorithm (GA), and Competitive Adaptive Resampling (CARS). Feature extraction included three spectral features (mean, standard deviation, and range) and five vegetation indices [Normalized Difference Vegetation Index (NDVI), Red Chlorophyll Index (RCI), Photochemical Reflectance Index (PRI), Anthocyanin Reflectance Index (ARI), and Plant Senescence Reflectance Index (PSRI)]. Each preprocessing-effective wavelength combination was applied to Support Vector Machine (SVM), Partial Least Squares Discriminant Analysis (PLS-DA), Random Forest (RF), and Extreme Gradient Boosting (XGBoost) models.

Results: Comparative analysis revealed that the SG-SPA-SVM and Raw-GA-SVM models achieved the highest classification performance, each with 97.8% accuracy on the test set. The proposed framework enabled early asymptomatic diagnosis as soon as 1 day post-inoculation, preceding visual symptom development.

Discussion: Integrating hyperspectral imaging with patch-based spatial-spectral analysis and machine learning enables accurate, early, and pathogen-specific detection of sesame leaf diseases. The reduced wavelength requirements of the GA-based approach improve feasibility for lightweight, cost-effective multispectral sensor development, thereby supporting sustainable sesame production and precision agriculture.

KEYWORDS

bacterial contamination, fungal contamination, hyperspectral imaging, machine learning, sesame

1 Introduction

Sesame (*Sesamum indicum* L.) is among the oldest oilseed crops cultivated by humans and holds considerable economic and nutritional importance worldwide (Namiki, 2007). Global sesame production is approximately 6 million tons annually, with major producing countries, such as Sudan, Myanmar, and India, accounting for more than 60% of the total production (Wei et al., 2022). Sesame is rich in protein (17.7 g/100 g), fat (55%), and essential minerals, including calcium, magnesium, and iron. Lignan compounds, including sesamin, sesamol, and sesamol, exhibit potent antioxidant, anti-inflammatory, and anticancer properties (Oboulbiga et al., 2023; Li et al., 2024). Additionally, sesame contains high levels of unsaturated fatty acids, with oleic acid accounting for 38.84% and linoleic acid for 46.26% of total fatty acids. Clinical studies indicate that sesame oil consumption reduces LDL cholesterol levels by 9.5% and increases HDL cholesterol levels, thereby offering benefits for cardiovascular disease prevention (Namayandeh et al., 2013; Aluganti Narasimhulu et al., 2018). However, this economically important crop is severely threatened by leaf blight caused by *Corynespora cassiicola* and bacterial leaf spot caused by *Pseudomonas syringae* both of which result in substantial damage to sesame cultivation. *Corynespora cassiicola* is a polyphagous pathogen infecting more than 530 plant species and prevalent in tropical and subtropical regions, causing yield losses of 10–42% in soybeans and considerable economic damage to cotton (Dixon et al., 2009; Sumabat et al., 2018; Edwards Molina et al., 2022). In Florida's tomato industry, this pathogen causes losses exceeding \$3,000 per acre within a market valued at \$500 million annually, demonstrating its destructive potential (MacKenzie et al., 2018). In sesame, the pathogen induces dark brown lesions and chlorosis in leaf tissues, causes premature defoliation and plant death, and severely reduces yield and quality (Jia et al., 2021). *Pseudomonas syringae* produces small water-soaked lesions on leaves that progress to total leaf necrosis. Although primarily reported in Asia, recent occurrences have also been documented in North America (Gardan et al., 1999; Firdous et al., 2009; Duin et al., 2024).

Existing disease detection methods rely on traditional techniques, which exhibit severe limitations. Visual detection requires experienced specialists, is labor-intensive, and depends on subjective judgment, leading to considerable inter-individual variability (Del Ponte et al., 2017; Bock et al., 2020). Laboratory-based detection offers high accuracy; however, it requires destructive sampling, prolonged analysis times, expensive equipment, and skilled operators, thereby presenting economic barriers to large-scale farm implementation (Law et al., 2014; Zhao et al., 2014). The primary limitation of existing methods is their ability to detect diseases solely after visible symptoms appear, resulting in substantial economic losses and missed early intervention opportunities (Mahlein, 2016; Lowe et al., 2017). Hyperspectral imaging (HSI) has emerged as a revolutionary alternative for overcoming these limitations. HSI is a non-destructive method that collects continuous spectral data for each pixel, resulting in three-dimensional data cubes capable of detecting subtle physiological and biochemical changes not visible to the naked eye. The applicable spectral range varies according to the sensor type:

visible–near infrared (VNIR) systems typically cover 400–1,000 nm, whereas short-wave infrared (SWIR) and extended systems can reach up to 2,500 nm. The primary benefit of HSI lies in its ability to identify diseases during asymptomatic stages by detecting changes in chlorophyll content, cellular structure, and water content several days to weeks before visible symptoms emerge (Lu et al., 2020; Nguyen et al., 2021). Recent agricultural studies using HSI for disease diagnosis have yielded promising results, with classification accuracies of 95.73% for soybean charcoal rot using three-dimensional convolutional neural networks (3D-CNNs), 98.5–99.9% for sugar beet *Cercospora*, and >98% for early detection of barley *Magnaporthe* within 24 h post-inoculation. This highlights its potential for asymptomatic early-stage infection detection (Arens et al., 2016; Zhou et al., 2018; Nagasubramanian et al., 2019).

Selecting key wavelengths critical for disease detection from numerous bands in hyperspectral data and applying them to machine learning (ML) algorithms is central to developing practical diagnostic systems. HSI consists of high-dimensional data containing hundreds of consecutive spectral bands ranging from 400 to 2,500 nm. Dimensionality reduction is essential because of the high correlation and redundancy between adjacent wavelengths (Lowe et al., 2017; Fabelo et al., 2018). Effective wavelength selection enhances processing speed by reducing computational complexity, improves model performance by minimizing noise, identifies physiologically relevant wavelengths, and provides a basis for practical multispectral sensor design (Elmasry et al., 2012). Wavelength selection methods encompass the Successive Projection Algorithm (SPA), the Genetic Algorithm (GA), and Competitive Adaptive Reweighted Sampling (CARS). Studies have demonstrated that selecting sets of 10–30 wavelengths can achieve classification performance equivalent to that of using the full spectrum (Li et al., 2009; Zhu et al., 2017). Various ML algorithms have been applied to plant disease detection, each with unique characteristics. Support Vector Machine (SVM) is particularly effective for high-dimensional datasets, particularly when the feature count exceeds the sample size. Using an RBF kernel for nonlinear separation demonstrates excellent performance, achieving 97% accuracy in sugar beet disease detection (Hearst et al., 1998; Rumpf et al., 2010; Mahlein et al., 2013). Partial Least Squares Discriminant Analysis (PLS-DA) is a supervised learning method effective for class discrimination in high-dimensional hyperspectral data. It is particularly suited for classification and biomarker selection, demonstrating enhanced performance in metabolomics and plant trait prediction studies. This method offers simultaneous dimensionality reduction and discriminant information preservation (Barker and Rayens, 2003; Burnett et al., 2021). Random Forest (RF) is an ensemble method of multiple decision trees. It demonstrates resilience to overfitting, facilitates variable importance assessment for high interpretability, and offers the advantage of relatively straightforward hyperparameter tuning (Breiman, 2001; Szymańska et al., 2012; Belgiu and Drăguț, 2016). eXtreme Gradient Boosting (XGBoost) is an advanced gradient boosting algorithm that achieves high predictive performance via sequential weak learner ensembles. It demonstrates a coefficient of determination (R^2) = 0.85 accuracy in

HSI-based plant metabolite prediction and 89% classification accuracy in wheat fusarium head blight detection (Chen and Guestrin, 2016; Yoon et al., 2023; Mustafa et al., 2024).

This study introduces a 3×3 -pixel patch-based analysis aimed at optimizing computational efficiency while effectively utilizing spatial information. Small spatial windows effectively capture fine features of lesion boundaries while maintaining low computational complexity (Mohanty et al., 2016). The perilesional zone, located adjacent to healthy tissue, is particularly important because it represents an active region of pathogen spread and plant defense responses, offering essential insights for early detection (Kuska et al., 2015). By integrating the spatial relationships between the central pixel and its eight neighboring pixels, this approach improves robustness to sensor noise and achieves higher classification accuracy than single-pixel analysis (Qin et al., 2025). This spatial-spectral integration approach represents a balanced method that effectively leverages the high-dimensional data of hyperspectral imagery while mitigating the risk of overfitting (Nagasubramanian et al., 2019).

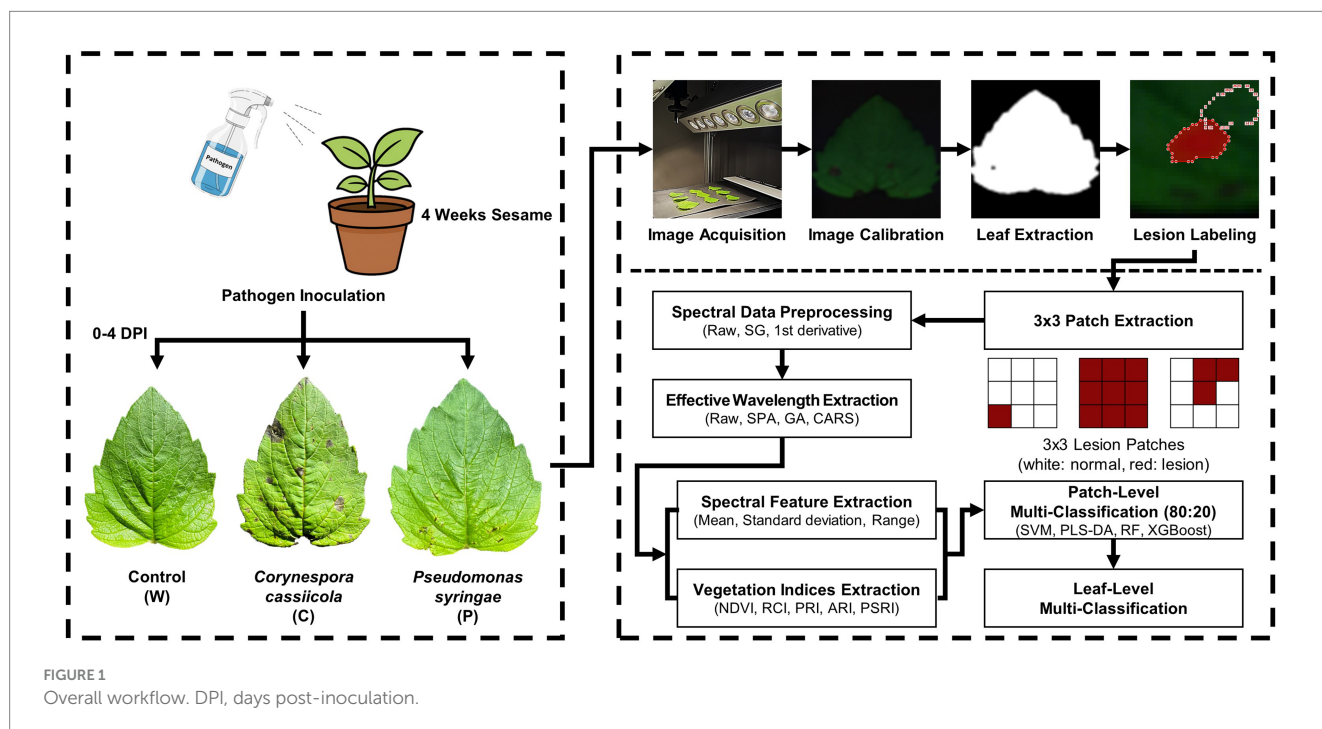
To date, existing research on sesame disease detection have been extremely limited, with no studies addressing multi-pathogen classification via integrated spatial-spectral analysis (Grieve et al., 2019; Ram et al., 2024). To address this gap, this study aimed to develop a comprehensive framework for the early diagnosis and multi-classification of sesame leaf diseases using HSI and ML. The specific objectives were as follows: (1) to analyze the spectral differences among control, *C. cassiicola*, and *P. syringae* infections during the early post-inoculation stage (0–4 days); (2) to implement a 3×3 patch-based spatial-spectral fusion to capture disease progression patterns; (3) to evaluate combinations of three preprocessing methods [Raw reflectance spectra, Savitzky–Golay (SG) smoothing, and 1st derivative] and four wavelength selection algorithms (Raw, SPA, GA, and CARS) for identifying optimal spectral features; (4) to integrate three spectral features, including mean, standard deviation, range, with five complementary vegetation indices, including the Normalized Difference Vegetation Index (NDVI), Red

Chlorophyll Index (RCI), Photochemical Reflectance Index (PRI), Anthocyanin Reflectance Index (ARI), Plant Senescence Reflectance Index (PSRI), in quantifying diverse physiological responses to pathogen infection; (5) to compare and evaluate four ML algorithms (SVM, PLS-DA, RF, and XGBoost) via systematic evaluation of multi-class classification tasks. The proposed methodology is anticipated to achieve robust early-stage disease detection by integrating innovative patch-based analysis, comprehensive spectral data preprocessing, and ML approaches. This integrated framework is expected to not only advance the fundamental understanding of host-pathogen interactions via hyperspectral signatures but also offers a practical solution for real-time disease monitoring in sesame production systems.

2 Materials and methods

2.1 Overall workflow

The overall workflow for the early detection of sesame leaf diseases is presented in Figure 1. The process consisted of the following steps: (a) pathogen inoculation of sesame plants with control (W), *C. cassiicola* (C), and *P. syringae* (P) treatments at 0–4 days post-inoculation (DPI); (b) hyperspectral image acquisition using SPECIM FX10e sensor; (c) image calibration with white/dark references; (d) leaf extraction via NDVI-based segmentation; (e) lesion labeling for supervised learning; (f) spectral data preprocessing using Raw, SG, and 1st derivative methods; (g) 3×3 patch extraction from labeled regions; (h) effective wavelength extraction using Raw spectrum, SPA, GA, and CARS algorithms; (i) spectral data feature extraction, encompassing statistical measures (mean, standard deviation, and range) and vegetation indices (NDVI, RCI, PRI, ARI, and PSRI); (j) patch-level multi-classification using SVM, PLS-DA, RF, and XGBoost models; and (k) leaf-level classification to facilitate final disease diagnosis.



2.2 Plant materials and experimental design

The sesame (*Sesamum indicum* L.) cultivar ‘Haniol’ was selected and sourced from the Rural Development Administration of Korea due to its commercial importance and susceptibility to leaf diseases. Seeds were disinfected with benzimidazole-based fungicide, washed three times with sterile water, and sown in plastic pots (diameter 120 mm, height 115 mm) filled with commercial potting soil for cultivation in a plant growth chamber. Plants were grown under conditions of 28/25 °C (day/night), 30–80% relative humidity, 16/8-h photoperiod, and 250 $\mu\text{mol m}^{-2} \text{s}^{-1}$ photosynthetic photon flux density. Four weeks post-sowing, uniform plants at the V3 growth stage were selected and the third true leaf was used for the experiments. A total of 255 individuals were used in this study and divided into three treatment groups: (1) sterile water inoculation control group, W; (2) *C. cassiicola* MR 4431 inoculation group, C; and (3) *P. syringae* MR 45261 inoculation group, P. Each treatment group comprised 85 individuals, and experiments were performed in triplicate. Leaves were harvested from pot-grown plants, and hyperspectral images were captured at 0 DPI. The leaves were subsequently spray-inoculated with pathogens and transferred to plastic containers (295 × 230 × 85 mm) to maintain high humidity.

2.3 Pathogen culture and inoculation

The *C. cassiicola* MR 4431 and *P. syringae* MR 45261 strains were obtained from the Smart Agricultural Technology Division (National Institute of Crop Science, Rural Development Administration of the Republic of Korea). The *C. cassiicola* MR 4431 strain was cultured on Potato Dextrose Agar (PDA) medium at 25 °C in the dark for 7 days. Spores were harvested by washing the culture surface with sterile water and filtering it through a gauze. The spore concentration was adjusted to 1×10^5 spores mL^{-1} using a hemocytometer. The *P. syringae* MR 45261 strain was cultured on Tryptic Soy Agar (TSA) medium at 28 °C for 48 h. Bacterial cells were suspended in sterile water and adjusted to $\text{OD}_{600} = 1.0$ (approximately 1×10^9 cells mL^{-1}) using a spectrophotometer (DS-11+, DeNovix Inc., USA). The harvested leaves were placed in sterile plastic containers lined with sterile filter paper and moistened with sterile water to maintain high humidity. Inoculation was performed using a portable sprayer, and the containers were sealed and maintained in a growth chamber. Disease progression was monitored daily for four DPI (Figure 1).

2.4 Hyperspectral image acquisition

Hyperspectral images were acquired using a SPECIM FX10e sensor (Specim Ltd., Finland), covering a spectral range of 400–1,000 nm across 224 consecutive bands. The imaging system was configured in line-scanning mode with 1,232 spatial pixels, mounted 30 cm above the sample stage in a darkroom. The conveyor belt speed was optimized based on the working distance to achieve optimal spatial resolution. Illumination was provided by two halogen lamps positioned at 45° angles to minimize specular reflection. The hyperspectral imaging system is presented in Supplementary Figure 1. Imaging was conducted at the same time each day to minimize diurnal variations, and 150 hyperspectral images were acquired from 0 to 4 DPI. Radiometric calibration was performed using a 99% white reference panel (Spectralon, Labsphere Inc., USA), while dark current measurements were acquired

before each imaging session. The raw digital values were converted to reflectance values using the following Equation 1:

$$R = \frac{I_o - I_d}{I_w - I_d} \quad (1)$$

where R is the calibration image; I_o is the original image; I_d is the dark reference image; and I_w is the white reference image. Spectrally and radiometrically calibrated images were subsequently obtained.

2.5 Hyperspectral image processing

Hyperspectral images were analyzed using the NDVI for data segmentation and extraction of pure leaf areas. Subsequently, lesions within the leaf regions were manually labeled to generate region of interest (ROI) labels for supervised learning. The labeling timing was refined based on lesion development patterns, with lesion areas from samples at 4 DPI labeled for *C. cassiicola* MR 4431 and *P. syringae* MR 45261 when the symptoms were clearly observable. Healthy control samples (non-treated) were labeled by uniformly selecting data from 0 to 4 DPI. Spectral data were extracted from the original hyperspectral image data using ROI coordinates. Data extraction was performed for 224 bands within the 400–1,000 nm range; however, only 184 bands in the 450–950 nm range were used for data processing to remove noise.

A 3×3 -pixel patch-based approach was implemented to incorporate the spatial context. Spatial patches containing the central pixel and its eight neighboring pixels were extracted for each labeled pixel located at the coordinates (x, y) . Given a hyperspectral image I with dimensions $H \times W \times B$ (height, width, and bands), where each pixel contains a spectral vector, the 3×3 patch centered at positions $(x$ and $y)$ can be formally defined as follows (Equation 2):

$$\text{Patch}(x, y) = \{I_{i,\lambda} \mid i \in [x-1, x+1], j \in [y-1, y+1]\} \quad (2)$$

where $\text{Patch}(x, y)$ is the 3×3 patch centered at pixel (x, y) ; $I_{i,j}$ is the spectral vector at pixel position (i, j) ; (x, y) is the coordinates of the center pixel. Each patch retained complete spectral information for the center coordinates and number of included pixels (maximum 9).

The class label for each patch was assigned based on the annotation of the central pixel. Consequently, patches extracted from lesion boundary regions may comprise both diseased and healthy pixels. This design was intentional, as the perilesional zone (where healthy and diseased tissues interface) serves as an active region for pathogen spread and plant defense responses, thereby providing critical spectral signatures for early disease detection (Kuska et al., 2015).

Three spectral data preprocessing methods were subsequently evaluated from the patch data: (1) Raw reflectance spectra maintaining the original reflectance values; (2) SG for noise reduction (window = 11, polyorder = 2); and (3) 1st derivative to enhance absorption features and minimize baseline effects.

Following patch data preprocessing, four effective wavelength extraction algorithms were applied: (1) the raw spectrum includes all 184 bands; (2) the Successive Projection Algorithm (SPA) minimizes orthogonality with previously selected wavelengths to minimize redundant information; (3) the GA is an evolutionary algorithm

mimicking natural selection principles that searches for optimal wavelength combinations through crossover and mutation; (4) CARS iteratively assesses variable importance based on PLS regression coefficients. A grid search was employed across all methods to select 15–30 optimal effective wavelengths.

Three spectral feature measures and five vegetation indices were calculated for each wavelength from the patch data and employed as patch features (Table 1). Spectral features included the mean, standard deviation, and range of wavelength bands within the 3×3 patches. The vegetation indices used were the NDVI, RCI, PRI, ARI, and PSRI. The NDVI is the most widely used vegetation index that evaluates plant photosynthetic activity and biomass by normalizing the difference between red and near-infrared reflectance (Rouse et al., 1974). Healthy plants exhibit high values owing to red light absorption by chlorophyll and strong near-infrared reflection by cellular structures, whereas diseased plants present lower values. The RCI directly estimates chlorophyll content and can detect early chlorophyll degradation due to pathogen infection (Clevers and Gitelson, 2013). The PRI assesses photosynthetic light-use efficiency by detecting minor variations in wavelengths associated with the xanthophyll cycle (Gamon et al., 1997). It responds sensitively to short-term environmental stress, enabling the capture of physiological changes that arise during the early stages of pathogen infection. The ARI quantifies the anthocyanin pigments that accumulate in response to stress conditions (Gitelson et al., 2001). The PSRI assesses the chlorophyll-to-carotenoid ratio as a measure of plant senescence and stress levels; as disease progresses, chlorophyll degrades, while carotenoids remain relatively stable, potentially causing PSRI to increase (Merzlyak et al., 1999).

TABLE 1 Spectral features and vegetation indices.

Feature	Index	Formula
Spectrum	Mean	$\bar{R}_\lambda = \frac{1}{9} \sum_{i=1}^9 R_{i,\lambda}$
	Standard deviation	$\sigma_\lambda = \sqrt{\frac{1}{9} \sum_{i=1}^9 (R_{i,\lambda} - \bar{R}_\lambda)^2}$
	Range	$Range_\lambda = \max(R_{i,\lambda}) - \min(R_{i,\lambda})$
VIs	NDVI	$\frac{R_{800} - R_{670}}{R_{800} + R_{670}}$
	RCI	$\frac{R_{800}}{R_{670}} - 1$
	PRI	$\frac{R_{531} - R_{670}}{R_{531} + R_{670}}$
	ARI	$\frac{1}{R_{550}} - \frac{1}{R_{700}}$
	PSRI	$\frac{R_{678} - R_{500}}{R_{750}}$

2.6 Machine learning classification and visualization

Four ML algorithms (i.e., SVM, PLS-DA, RF, and XGBoost) were used to compare the multi-class classification performance among the control, *C. cassicola* MR 4431, and *P. syringae* MR 45261 treatment groups. The patch-level dataset was divided into training (80%) and test (20%) sets via stratified sampling. The leaf-level dataset was constructed using 27 newly collected leaves. Class weights were applied to the training data to address class-imbalance issues. The sample distribution was W:C:p = 3:1:1, with the control class (W) intentionally oversampled to capture a comprehensive range of normal spectral variability across diverse healthy tissue types. To address this class imbalance, inverse-ratio class weights (W:C:p = 1:3:3) were incorporated during model training. The hyperparameters for each algorithm were optimized using five-fold cross-validation and grid search methods to investigate the optimal parameter combinations. The SVM was configured with the regularization parameter C [1, 10], Radial Basis Function (RBF), and gamma parameters [scale, 0.01, 0.001]. The PLS-DA was set using 10–25 latent variables. The RF was configured with number of estimators [200, 300], maximum tree depth [10, 20, 30], and minimum sample number [5, 10]. XGBoost was set using boosting rounds [500], maximum tree depth [5, 7], learning rate [0.1, 0.01], subsample ratio [0.6, 0.8], and early stopping rounds [20] to prevent overfitting. The model performance was evaluated using an independent test set, employing various evaluation metrics, including accuracy, precision, recall, and F1-score Equations 3–6.

$$Accuracy = \frac{(TP + TN)}{(TP + TN + FP + FN)} \quad (3)$$

$$Precision = \frac{(TP)}{(TP + FP)} \quad (4)$$

$$Recall = \frac{(TP)}{(TP + FN)} \quad (5)$$

$$F1-score = 2 \times \frac{(Precision \times Recall)}{(Precision + Recall)} \quad (6)$$

TP, TN, FP, and FN represent true positives, true negatives, false positives, and false negatives, respectively.

Classification model development was performed using a hierarchical approach at both patch and leaf levels. In patch-level classification, the spectral features [mean, standard deviation, and range] and vegetation indices [NDVI, RCI, PRI, ARI, and PSRI] of each patch were used as feature vectors to classify the control (W), *C. cassicola* MR 4431 (C), and *P. syringae* MR 45261 (P) infection classes. Patch-level classification is designed to detect localized physiological changes caused by pathogen infections. In leaf-level classification, patches were generated from all pixels within individual leaf regions of the test set, the spectral features and vegetation indices of the patches were calculated, and the optimal patch-level multi-class classification model was applied using a lesion area ratio-based methodology. Aggregation of patch-wise classification results involved calculating the proportion of patches

classified as control, *C. cassiicola*, or *P. syringae* within each leaf region (Equations 7–9).

$$W \text{ rate (control\%)} = \frac{(W \text{ patches})}{(\text{all patches})} \times 100 \quad (7)$$

$$C \text{ rate (C.cassiicola\%)} = \frac{(C \text{ patches})}{(\text{all patches})} \times 100 \quad (8)$$

$$P \text{ rate (P.syringae\%)} = \frac{(P \text{ patches})}{(\text{all patches})} \times 100 \quad (9)$$

In the final diagnosis, Receiver Operating Characteristic (ROC)-based thresholding was applied, whereby calculated lesion area ratio values were compared with optimal thresholds derived from ROC analysis to determine final leaf classification. Comprehensive ROC curves was performed using a conservative true negative (TN) = 0 criterion (specificity = 100%) to ensure that all control samples were accurately classified as healthy controls.

2.7 Statistical analysis and data processing

All experiments were performed in triplicate, and the results are presented as mean $\pm$ standard deviation. Statistical significance among treatment groups was verified using one-way analysis of variance (ANOVA) and Tukey's HSD *post-hoc* test ($p < 0.05$). Duncan's multiple range test was used for multiple comparisons involving more than three groups to identify homogeneous subsets and rank the differences between treatments. Spectral data analysis, ML model construction, and statistical analyses were performed in Python 3.10 environment using scikit-learn, numpy, pandas, matplotlib, and seaborn libraries. The overall data processing workflow proceeded as follows: image acquisition and preprocessing, spectral extraction and labeling, patch generation, preprocessing application, effective wavelength selection, patch feature extraction, model training and validation, and performance evaluation. A systematic comparative analysis was conducted on 144 combinations (3 treatments $\times$ 3 preprocessing methods $\times$ 4 effective wavelength groups $\times$ 4 classification algorithms).

3 Results

3.1 Sesame leaf spectral characteristics analysis and preprocessing

Analysis of patch-level average spectral reflectance changes according to treatment status in sesame leaves [control (W), *C. cassiicola* MR 4431 (C), and *P. syringae* MR 45261 (P)] revealed distinct differences across the 450–950 nm wavelength range (Figure 2A). Particularly significant differences were observed in the visible light (520–680 nm) and near-infrared (720–800 nm) regions. The W leaves exhibited strong absorption peaks at approximately 680 nm owing to chlorophyll absorption and high reflectance in the near-infrared region above 720 nm. The C lesions showed the lowest reflectance values at 550 nm, followed by an increase in reflectance from 550 to 680 nm. This suggests chlorophyll degradation and a relative increase in carotenoid pigment levels. The P lesions consistently demonstrated high reflectance values in the visible light region of 500–700 nm, with a decreasing trend observed after 700 nm. A comparison of the three spectral data preprocessing methods revealed that SG smoothing was effective for noise reduction, whereas the 1st derivative emphasized subtle spectral differences. Rapid changes were detected at 520, 570, and 720 nm (Figures 2B,C). Spectral analysis revealed distinct reflectance patterns among the three classes (Figure 2), forming the basis for subsequent wavelength selection.

3.2 Effective wavelength extraction

After generating the 3×3 patches and completing each spectral data preprocessing step, four effective wavelength extraction algorithms (Raw, SPA, GA, and CARS) were applied to each preprocessing method, resulting in 12 preprocessing-effective wavelength selection combinations (Table 2, Supplementary Table 1). SPA selected the highest number of effective wavelengths, with 36, 39, and 38 wavelengths for the Raw, SG, and 1st derivative, respectively. Conversely, GA selected 32, 32, and 31 wavelengths, while CARS selected the fewest effective wavelengths, with 18, 17, and 16 wavelengths, respectively. SPA combinations (blue) demonstrated a broader distribution concentrated in the 450–550 nm region, whereas the GA and CARS combinations (green and red) exhibited relatively uniform distributions across the spectrum; however, they were clustered in the 500–600 nm and

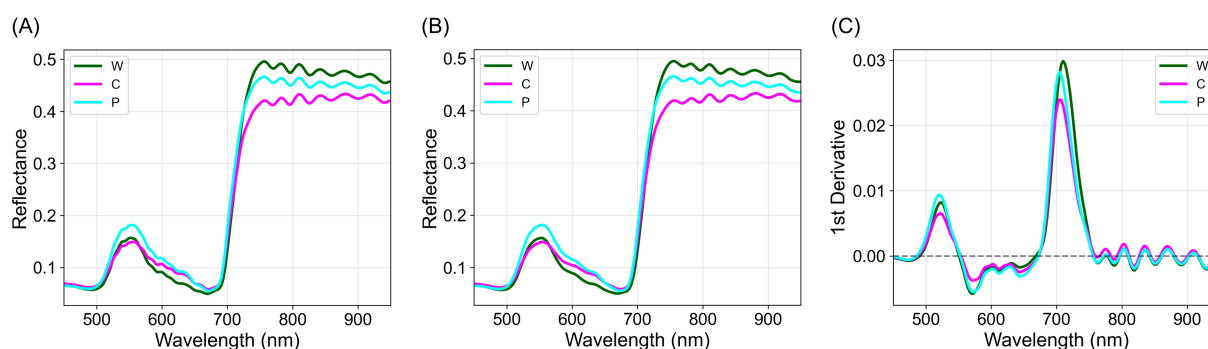


FIGURE 2

Average reflectance values for W-C-P classes after spectral data preprocessing. (A) Raw; (B) SG; (C) 1st derivative.

650–730 nm regions (Figure 3). This pattern is consistent with the differences identified in the spectral characteristic analysis, thereby confirming that numerous wavelengths were selected in these regions.

3.3 3×3 patch-based features extraction

Spatial-spectral integration analysis through 3 × 3 patch generation derived comprehensive feature sets from preprocessed spectral data and extracted spectral features and vegetation indices from each patch. The Raw-GA combination was randomly selected as representative (Figure 4). Mean spectral reflectance values exhibited significant gaps between the control and lesion patches in the 532–658 nm range, with the control patches showing relatively lower reflectance values. From 720 to 737 nm, the control reflectance values exceeded those of both lesion patches, maintaining the order of the control, C lesion, and P lesion up to 950 nm. This suggests heterogeneous tissue damage and irregular pigment distribution within the 3 × 3 patch area. Spectral standard deviation and range showed minimal variability in the control patches, whereas both lesion patches demonstrated sharp differences from the control in the 532–658 nm range, increasing from 699 nm and maintaining approximately two-fold differences beyond 720 nm. This indicates substantial spatial heterogeneity, where adjacent pixels within the same patch show markedly different spectral responses owing to localized disease progression patterns in the region beyond 700 nm.

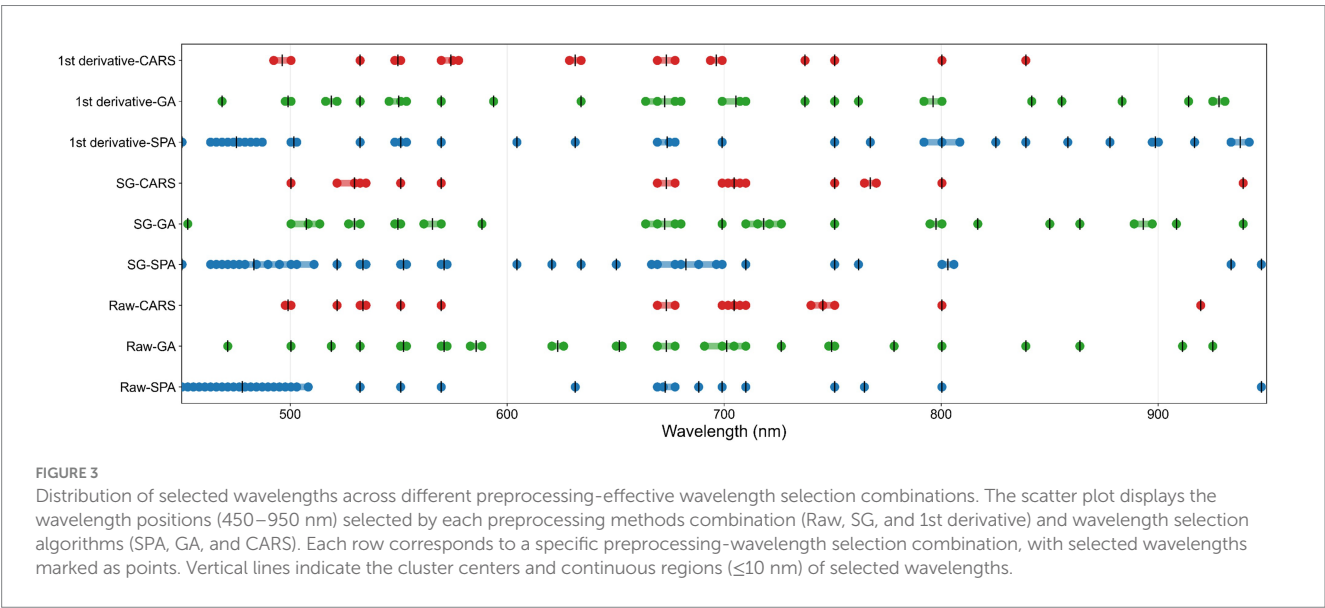
TABLE 2 Number of effective wavelengths across preprocessing methods.

Preprocessing			
Effective wavelength	Raw	SG	1st derivative
Raw	184	184	184
SPA	36	39	38
GA	29	30	30
CARS	19	19	18

The vegetation index analysis revealed distinct physiological characteristics among the treatment groups. NDVI was highest for control patches ($\mu = 0.815$), followed by P lesions ($\mu = 0.782$), and lowest for C lesions ($\mu = 0.761$) (Figure 5). This reflects a progressive reduction in photosynthetic activity and chlorophyll content. RCI demonstrated the clearest separation among treatment states, with control patches ($\mu = 9.081$) maintaining significantly higher values than those of both C and P lesion patches (C: $\mu = 6.674$, P: $\mu = 7.381$), thereby confirming substantial chlorophyll degradation in both lesions regardless of pathogen type. PRI revealed differential stress responses between the two lesion patches: control patches showed positive values ($\mu = 0.010$), whereas C and P lesions exhibited negative values (C: $\mu = -0.028$, P: $\mu = -0.014$). This indicates photosynthetic stress induction. ARI showed an opposite pattern to that of other vegetation indices, with control patches exhibiting the lowest values ($\mu = -1.896$) and both C and P lesion patches showing significant increases (C: $\mu = -0.733$, P: $\mu = -0.035$). This suggests that anthocyanin accumulation is a universal stress response mechanism across the C- and P-lesion patch types. PSRI showed no significant differences among the control, C, and P lesions, thereby suggesting that it may represent a mechanism distinct from senescence processes.

3.4 Patch-level multi-classification model performance comparison

Four ML algorithms were applied using 12 preprocessing-effective wavelength group combinations for patch-level multi-classification of three classes (W, C, and P). The models with the highest F1 scores for each multi-classification algorithm are listed in Table 3. Among the SVM algorithms, the multi-classification models with the SG-SPA and Raw-GA treatment combinations achieved a test dataset accuracy of 0.978 and an F1 score of 0.978, thereby reflecting the highest performance among all preprocessing-effective wavelength group-classification model algorithm combinations. Among the PLS-DA and XGBoost algorithms, the multi-classification models with Raw-GA treatment combinations



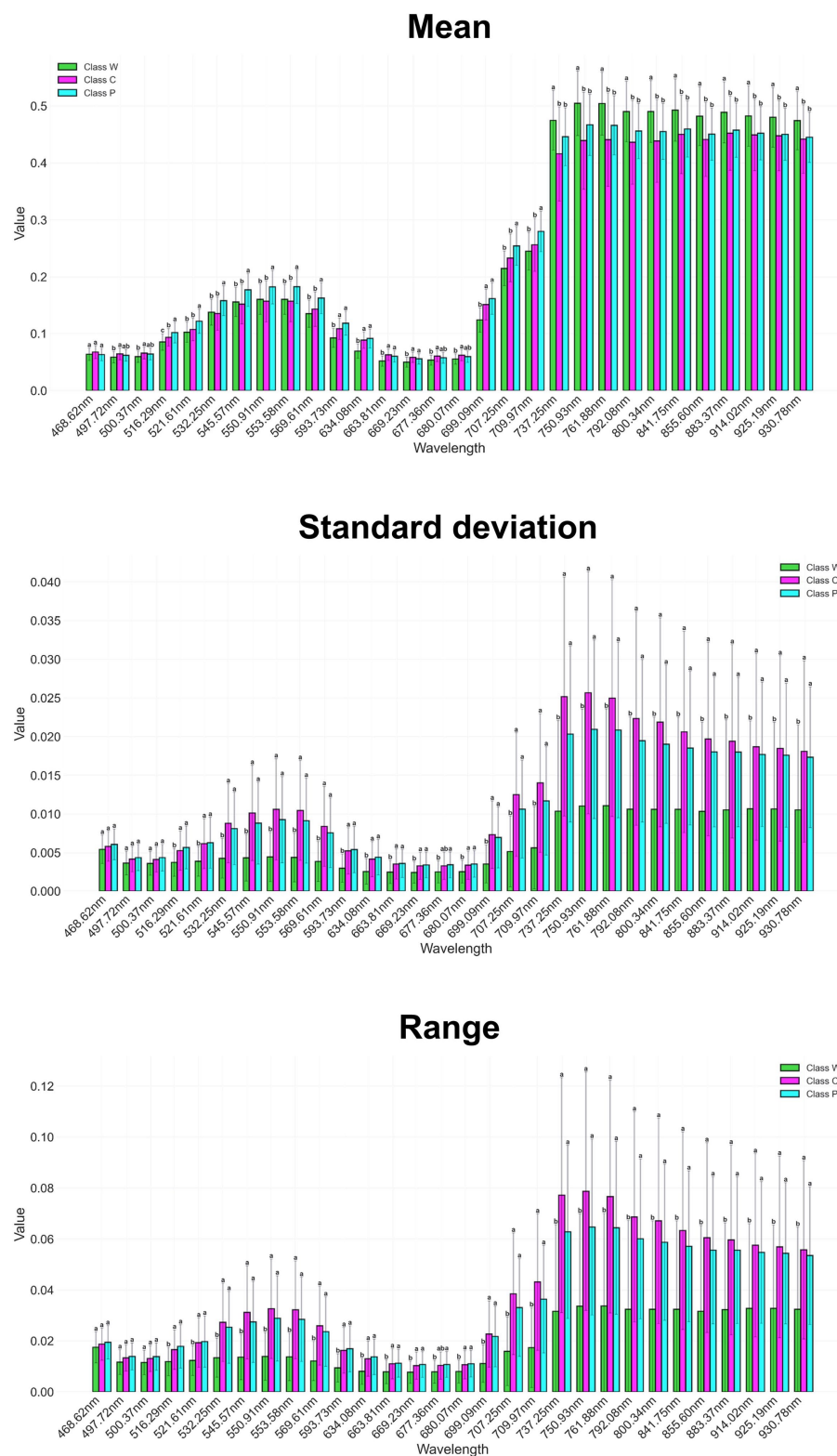


FIGURE 4

Spectral statistical features for W-C-P class patches of Raw-GA combination. Bar plots display reflectance, standard deviation, and range values for selected effective wavelengths. Error bars represent the standard deviation of the mean. Different lowercase letters above bars indicate statistically significant differences (Duncan's multiple range test, $p < 0.05$).

showed the highest performance, with accuracies of 0.844 and 0.965 and F1 scores of 0.837 and 0.965, respectively. Among the RF algorithms, the multi-classification model with the SG-SPA

treatment combination showed a test dataset accuracy of 0.919 and an F1 score of 0.918, followed by the Raw-CARS combination with an accuracy of 0.918 and an F1 score of 0.917. Overall, SVM and

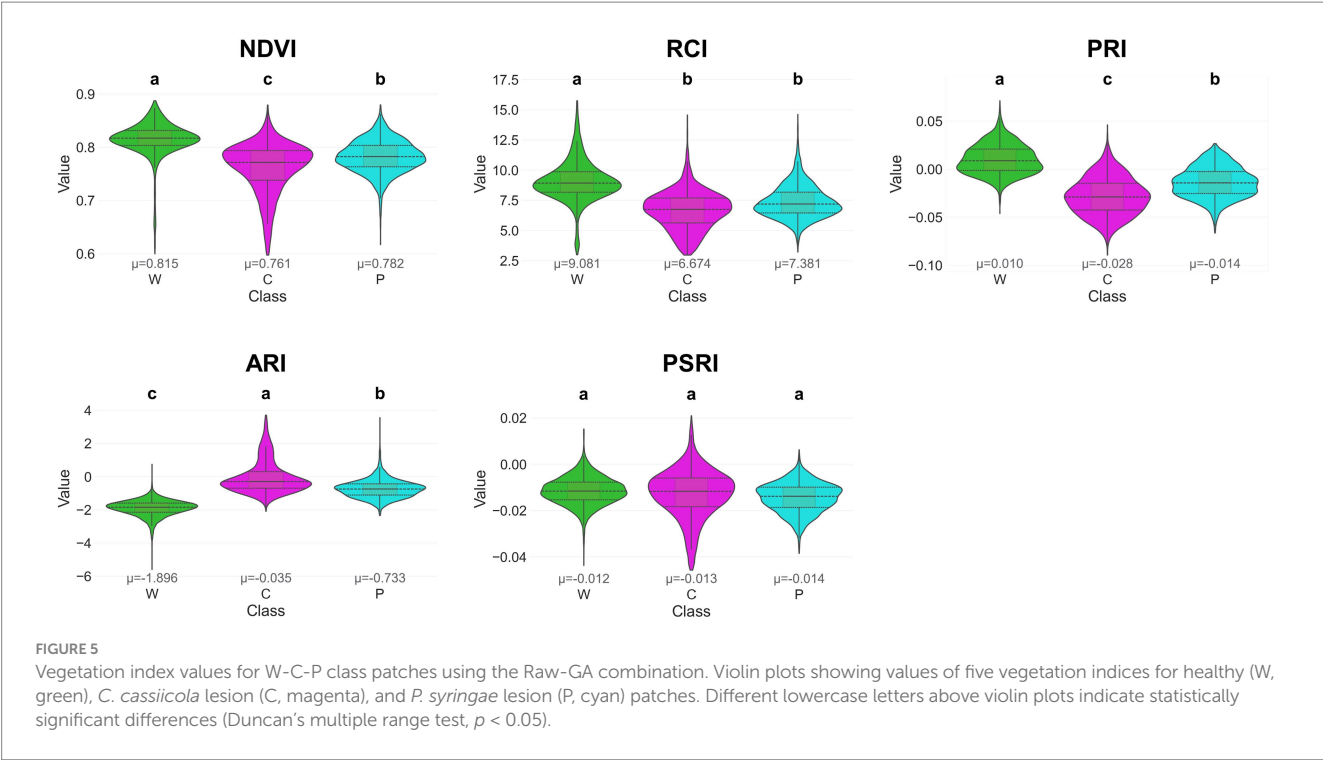


TABLE 3 Top five results per model for the 3 × 3 patch-level multi-classification models.

Model	Preprocessing	Effective wavelength (n)	Train accuracy	Test accuracy	F1-score
SVM	SG	SPA (39)	0.993	0.978	0.978
	Raw	GA (29)	0.991	0.978	0.978
	Raw	SPA (36)	0.997	0.972	0.972
	1st derivative	GA (30)	0.989	0.971	0.971
	SG	GA (30)	0.988	0.970	0.970
PLS-DA	Raw	GA (29)	0.925	0.844	0.837
	SG	SPA (39)	0.919	0.828	0.820
	1st derivative	GA (30)	0.916	0.822	0.813
	1st derivative	Raw (184)	0.912	0.817	0.810
	Raw	Raw (184)	0.912	0.817	0.810
RF	SG	SPA (39)	0.999	0.919	0.918
	Raw	CARS (19)	0.999	0.918	0.917
	Raw	GA (29)	0.999	0.915	0.914
	1st derivative	GA (30)	0.999	0.914	0.913
	Raw	SPA (36)	0.998	0.913	0.912
XGBoost	Raw	GA (29)	1	0.965	0.965
	SG	SPA (39)	1	0.962	0.962
	SG	GA (30)	1	0.962	0.962
	1st derivative	GA (30)	1	0.961	0.961
	Raw	SPA (36)	1	0.960	0.959

n, number of effective wavelengths.

XGBoost exhibited the highest classification performance, followed by RF, with PLS-DA showing the lowest classification performance. A detailed class-wise performance analysis of the optimal

SG-SPA-SVM and Raw-GA-SVM multi-classification models revealed that the control patches achieved a precision of 0.987, a recall of 0.992, and an F1 score of 0.989; the C lesion patches

achieved a precision of 0.974, a recall of 0.955, and an F1-score of 0.965; and the P lesion patches achieved a precision of 0.964, a recall of 0.978, and an F1-score of 0.971.

3.5 Patch-based leaf-level multi-classification for early detection performance evaluation

Building on these patch-level classification results, a hierarchical approach was employed to translate patch-level sensitivity for detecting localized physiological changes into practical leaf-level early diagnosis. The proportional distribution of patch classifications across each leaf determines the overall disease status through lesion area ratio analysis.

To evaluate the early diagnostic capability of the patch-level multi-classification models, a patch-to-leaf-level multi-classification approach was implemented at 0–3 DPI (Figure 6). Individual patch-based predictions applied the Raw-GA-SVM combination classification model, which exhibited the highest classification performance with relatively few effective wavelengths, to new test samples, where the proportional distribution of patches classified into each class determined the overall leaf condition. At 0 DPI, the entire sample population was classified as healthy W class, showing no immediately detectable spectral changes, and all individuals exhibited normal status. Additionally, all W-class individuals were accurately classified as W at

0–3 DPI. Disease diagnostic capability emerged at 1 DPI, with 8 and 17 individuals correctly classified as C or P, respectively (correct W: 26/26, C: 8/26, P: 17/26). At 2 DPI, C showed perfect classification before a visual diagnosis was possible, whereas P added one correctly classified individual at 1 DPI (correct W: 26/26, C: 26/26, P: 18/26). At 3 DPI, when subtle visual diagnosis was possible, accurate classification of all W and C samples and improved classification performance for P samples were confirmed (correct W: 26/26, C: 26/26, P: 23/26).

4 Discussion

This study successfully demonstrated the effectiveness of a patch-based spatial-spectral integration framework for the early diagnosis of sesame leaf diseases by evaluating preprocessing methods, wavelength selection algorithms, feature extraction, and ML models.

The selection of a 3×3 patch size for spatial-spectral integration was based on several considerations. First, the 3×3 configuration represents the minimum spatial unit that captures the central pixel with its complete first-order neighborhood, enabling effective characterization of the perilesional zone where pathogen spread and plant defense responses occur (Kuska et al., 2015). Second, small spatial windows effectively capture fine lesion boundary features while minimizing computational complexity (Mohanty et al., 2016). Third, larger patches (e.g., 5×5 or 7×7) would incorporate excessive

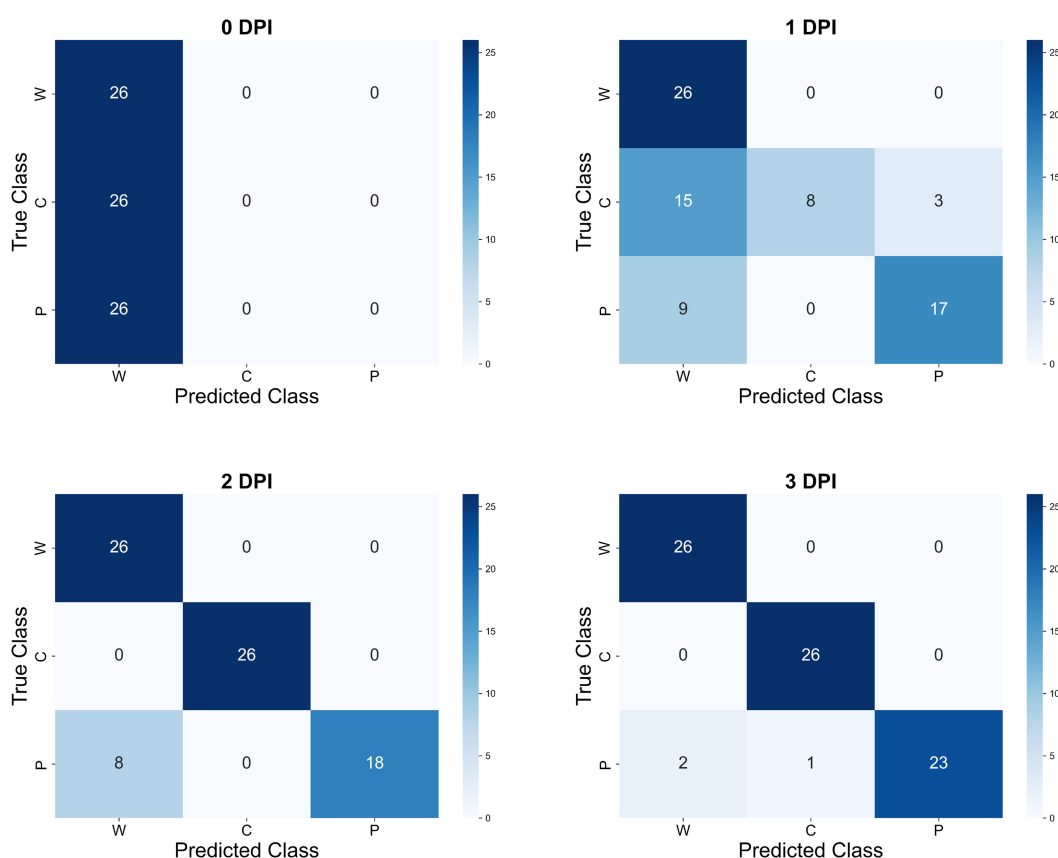


FIGURE 6

Confusion matrices illustrating classification performance of the Raw-GA-SVM model for the control (W), *C. cassiicola* (C) infected leaves, and *P. syringae* (P) infected leaves at 3 DPI.

heterogeneous tissue, potentially diluting disease-specific spectral signatures. Mahlein et al. (2012) demonstrated that as spatial resolution decreases, the number of mixed pixels increases and reflectance differences between healthy and diseased tissue diminish—a critical factor during early infection stages when lesions remain small. Finally, the 3×3 spatial configuration aligns with kernel dimensions commonly employed in spectral–spatial feature extraction for hyperspectral classification (Nagasubramanian et al., 2019).

Although both Raw-GA-SVM and SG-SPA-SVM achieved the same classification accuracy of 97.8% (Table 3), the GA-based approach was preferred because of its superior wavelength efficiency, employing 29 effective wavelengths compared with the 39 wavelengths in the SPA-based approach. Fewer spectral bands can substantially reduce sensor complexity, weight, and power consumption, all while maintaining diagnostic performance. This presents a clear advantage for the development of cost-effective multispectral sensors in drone-based applications (Govender et al., 2009; Aasen et al., 2018). The superior multi-classification performance of SVM stems from its suitability for high-dimensional spectral data. It effectively manages complex decision boundaries using RBF kernels and facilitates margin maximization, thereby enabling robust generalization despite overfitting resilience and spectral variations caused by environmental factors (Nagasubramanian et al., 2018).

The wavelengths depicted in Figure 3 correspond to physiologically important regions, specifically the green peak (500–600 nm) and red-edge (650–730 nm) regions, which illustrate the complex reflectance characteristics of carotenoids and chlorophyll (Lowe et al., 2017; Falcioni et al., 2023). These wavelength regions detect early photosynthetic pigment degradation and stress pigment accumulation due to pathogen infection. Distinct spectral signature differences between healthy and infected tissues were confirmed by observing increased reflectance values in the 500–600 nm region and decreased chlorophyll absorption peaks in the 650–730 nm region in *C. cassiicola* and *P. syringae* lesions (Figure 2) (Zhao et al., 2016).

Accurate multi-classification of fungal (*C. cassiicola*) and bacterial (*P. syringae*) pathogens enables the development of pathogen-specific control strategies, contributing to reduced pesticide use and environmental burden mitigation through targeted application of antifungal and antibacterial agents (Iost Filho et al., 2022; Dara et al., 2023). Pathogen-specific tailored control is expected to become a key element in sustainable agriculture by suppressing pesticide resistance and minimizing effects on non-target organisms (Morales et al., 2021). This study addresses an important gap in sesame disease research, as hyperspectral-based diagnostic approaches for sesame are lacking. The patch-based framework and early diagnostic capability offer substantial practical value for precision agriculture applications.

Data availability statement

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

Author contributions

T-AK: Data curation, Formal analysis, Writing – original draft. C-YA: Writing – review & editing. HJ: Writing – review & editing. Y-JP:

Writing – review & editing. JP: Writing – review & editing. JR: Writing – review & editing. HB: Writing – review & editing. YY: Writing – review & editing. HL: Writing – original draft, Writing – review & editing.

Funding

The author(s) declared that financial support was received for this work and/or its publication. This study was supported by the Fund of Science and Technology R&D Program of the Rural Development Administration (RDA), Republic of Korea (grant number PJ01740403).

Acknowledgments

This study was supported by 2025 collaborative research program between university and Rural Development Administration, Republic of Korea.

Conflict of interest

The author(s) declared that this work was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Generative AI statement

The author(s) declared that Generative AI was used in the creation of this manuscript.

Any alternative text (alt text) provided alongside figures in this article has been generated by Frontiers with the support of artificial intelligence and reasonable efforts have been made to ensure accuracy, including review by the authors wherever possible. If you identify any issues, please contact us.

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/fsufs.2025.1716907/full#supplementary-material>

SUPPLEMENTARY FIGURE 1

Hyperspectral imaging system setup.

SUPPLEMENTARY TABLE 1

Selected wavelengths across preprocessing–algorithm combinations.

References

- Aasen, H., Honkavaara, E., Lucieer, A., and Zarco-Tejada, P. J. (2018). Quantitative remote sensing at ultra-high resolution with UAV spectroscopy: a review of sensor technology, measurement procedures, and data correction workflows. *Remote Sens.* 10:1091. doi: 10.3390/rs10071091
- Aluganti Narasimhulu, C., Burge, K. Y., Doomra, M., Riad, A., and Parthasarathy, S. (2018). Primary prevention of atherosclerosis by pretreatment of low-density lipoprotein receptor knockout mice with sesame oil and its aqueous components. *Sci. Rep.* 8:12270. doi: 10.1038/s41598-018-29849-x
- Arens, N., Backhaus, A., Döll, S., Fischer, S., Seiffert, U., and Mock, H. P. (2016). Non-invasive presymptomatic detection of *Cercospora beticola* infection and identification of early metabolic responses in sugar beet. *Front. Plant Sci.* 7:1377. doi: 10.3389/fpls.2016.01377
- Barker, M., and Rayens, W. (2003). Partial least squares for discrimination. *J. Chemom.* 17, 166–173. doi: 10.1002/cem.785
- Belgiu, M., and Drăguț, L. (2016). Random forest in remote sensing: a review of applications and future directions. *ISPRS J. Photogramm. Remote Sens.* 114, 24–31. doi: 10.1016/j.isprsjprs.2016.01.011
- Bock, C. H., Barbedo, J. G. A., Del Ponte, E. M., Bohnenkamp, D., and Mahlein, A. K. (2020). From visual estimates to fully automated sensor-based measurements of plant disease severity: status and challenges for improving accuracy. *Phytopathol. Res.* 2:9. doi: 10.1186/s42483-020-00049-8
- Breiman, L. (2001). Random forests. *Mach. Learn.* 45, 5–32. doi: 10.1023/A:1010933404324
- Burnett, A. C., Anderson, J., Davidson, K. J., Ely, K. S., Lamour, J., Li, Q., et al. (2021). A best-practice guide to predicting plant traits from leaf-level hyperspectral data using partial least squares regression. *J. Exp. Bot.* 72, 6175–6189. doi: 10.1093/jxb/erab295
- Chen, T., and Guestrin, C. (2016). Xgboost: a scalable tree boosting system, in proceedings of the 22nd ACM SIGKDD international conference on knowledge discovery and data mining. New York, NY, USA: ACM, 785–794.
- Clevers, J. G. P. W., and Gitelson, A. A. (2013). Remote estimation of crop and grass chlorophyll and nitrogen content using red-edge bands on Sentinel-2 and -3. *Int. J. Appl. Earth Obs. Geoinf.* 23, 344–351. doi: 10.1016/j.jag.2012.10.008
- Dara, S. K., Rodriguez-Saona, C., and Morrison, W. R. III (2023). Editorial: integrated pest management strategies for sustainable food production. *Front. Sustain. Food Syst.* 7:1224604. doi: 10.3389/fsufs.2023.1224604
- Del Ponte, E. M., Pethybridge, S. J., Bock, C. H., Michereff, S. J., Machado, F. J., and Spolti, P. (2017). Standard area diagrams for aiding severity estimation: scientometrics, pathosystems, and methodological trends in the last 25 years. *Phytopathology* 107, 1161–1174. doi: 10.1094/PHYTO-02-17-0069-FI
- Dixon, L. J., Schlub, R. L., Pernezný, K., and Datnoff, L. E. (2009). Host specialization and phylogenetic diversity of *Corynespora cassiicola*. *Plant Pathol.* 99, 1015–1027. doi: 10.1094/PHYTO-99-9-1015
- Duin, I. M., Montalban, K. M., Joglekar, P., D'Amico-Willman, K., Ritchie, D. F., Fagen, J. R., et al. (2024). First report of bacterial leaf spot disease on sesame (*Sesamum indicum*) caused by *Pseudomonas syringae* pv. *Sesami* in North Carolina. *Plant Dis.* 108:1880. doi: 10.1094/PDIS-02-24-0292-PDN
- Edwards Molina, J. P., Navarro, B. L., Allen, T. W., and Godoy, C. V. (2022). Soybean target spot caused by *Corynespora cassiicola*: a resurgent disease in the Americas. *Trop. Plant Pathol.* 47, 315–331. doi: 10.1007/s40858-022-00495-z
- Elmasry, G., Kamruzzaman, M., Sun, D. W., and Allen, P. (2012). Principles and applications of hyperspectral imaging in quality evaluation of agro-food products: a review. *Crit. Rev. Food Sci. Nutr.* 52, 999–1023. doi: 10.1080/10408398.2010.543495
- Fabelo, H., Ortega, S., Ravi, D., Kiran, B. R., Sosa, C., Bulters, D., et al. (2018). Spatio-spectral classification of hyperspectral images for brain cancer detection during surgical operations. *PLoS One* 13:e0193721. doi: 10.1371/journal.pone.0193721
- Falcioni, R., Antunes, W. C., Demattê, J. A. M., and Nanni, M. R. (2023). A novel method for estimating chlorophyll and carotenoid concentrations in leaves: a two hyperspectral sensor approach. *Sensors* 23:3843. doi: 10.3390/s23083843
- Firdous, S. S., Asghar, R., Ul-Haque, M. I., Waheed, A., Afzal, S. N., and Mirza, M. Y. (2009). Pathogenesis of *Pseudomonas syringae* pv. *Sesami* associated with sesame (*Sesamum indicum* L.) bacterial leaf spot. *Pak. J. Bot.* 41, 927–934.
- Gamon, J. A., Serrano, L., and Surfus, J. S. (1997). The photochemical reflectance index: an optical indicator of photosynthetic radiation use efficiency across species, functional types, and nutrient levels. *Oecologia* 112, 492–501. doi: 10.1007/s004420050337
- Gardan, L., Shafik, H., Belouin, S., Broch, R., Grimont, F., and Grimont, P. A. D. (1999). DNA relatedness among the pathovars of *pseudomonas syringae* and description of *Pseudomonas tremiae* sp. nov. and *Pseudomonas cannabina* sp. nov. (ex Sute and Dowson 1959). *Int. J. Syst. Bacteriol.* 49, 469–478. doi: 10.1099/00207713-49-2-469
- Gitelson, A. A., Merzlyak, M. N., and Chivkunova, O. B. (2001). Optical properties and nondestructive estimation of anthocyanin content in plant leaves. *Photochem. Photobiol.* 74, 38–45. doi: 10.1562/0031-8655(2001)074<0038:opaneo>2.0.co;2
- Govender, M., Chetty, K., and Bulcock, H. (2009). A review of hyperspectral remote sensing and its application in vegetation and water resource studies. *Water SA* 33, 145–151. doi: 10.4314/wsa.v33i2.49049
- Grieve, B. D., Duckett, T., Collison, M., Boyd, L., West, J., Yin, H., et al. (2019). The challenges posed by global broadacre crops in delivering smart Agri-robotic solutions: a fundamental rethink is required. *Glob. Food Secur.* 23, 116–124. doi: 10.1016/j.gfs.2019.04.011
- Hearst, M. A., Dumais, S. T., Osuna, E., Platt, J., and Scholkopf, B. (1998). Support vector machines. *IEEE Intell. Syst. Appl.* 13, 18–28. doi: 10.1109/5254.708428
- Iost Filho, F. H., Pazini, J. B., Alves, T. M., Koch, R. L., and Yamamoto, P. T. (2022). How does the digital transformation of agriculture affect the implementation of integrated Pest management? *Front. Sustain. Food Syst.* 6:972213. doi: 10.3389/fsufs.2022.972213
- Jia, M., Liu, X., Zhao, H., Ni, Y., Liu, H., and Tian, B. (2021). Cell-wall-degrading enzymes produced by sesame leaf spot pathogen *Corynespora cassiicola*. *J. Phytopathol.* 169, 186–192. doi: 10.1111/jph.12973
- Kuska, M., Wahabzada, M., Leucker, M., Dehne, H. W., Kersting, K., Oerke, E. C., et al. (2015). Hyperspectral phenotyping on the microscopic scale: towards automated characterization of plant-pathogen interactions. *Plant Methods* 11:28. doi: 10.1186/s13007-015-0073-7
- Law, J. W. F., Ab Mutalib, N. S., Chan, K. G., and Lee, L. H. (2014). Rapid methods for the detection of foodborne bacterial pathogens: principles, applications, advantages and limitations. *Front. Microbiol.* 5:770. doi: 10.3389/fmicb.2014.00770
- 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, M., Luo, J., Nawaz, M. A., Stockmann, R., Buckow, R., Barrow, C., et al. (2024). Phytochemistry, bioaccessibility, and bioactivities of sesame seeds: an overview. *Food Rev. Int.* 40, 309–335. doi: 10.1080/87559129.2023.2168280
- 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
- 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
- MacKenzie, K. J., Sumabat, L. G., Xavier, K. V., and Vallad, G. E. (2018). A review of *Corynespora cassiicola* and its increasing relevance to tomato in Florida. *Plant Health Prog.* 19, 303–309. doi: 10.1094/PHP-05-18-0023-RV
- Mahlein, A. K. (2016). Plant disease detection by imaging sensors—parallels and specific demands for precision agriculture and plant phenotyping. *Plant Dis.* 100, 241–251. doi: 10.1094/PDIS-03-15-0340-FE
- Mahlein, A. K., Rumpf, T., Welke, P., Dehne, H. W., Plümer, L., Steiner, U., et al. (2013). Development of spectral indices for detecting and identifying plant diseases. *Remote Sens. Environ.* 128, 21–30. doi: 10.1016/j.rse.2012.09.019
- Mahlein, A. K., Rumpf, T., Welke, P., Dehne, H. W., Plümer, L., Steiner, U., et al. (2012). Hyperspectral imaging for small-scale analysis of symptoms caused by different sugar beet diseases. *Plant Methods* 8:3. doi: 10.1186/1746-4811-8-3
- Merzlyak, M. N., Gitelson, A. A., Chivkunova, O. B., and Rakitin, V. Y. (1999). Non-destructive optical detection of pigment changes during leaf senescence and fruit ripening. *Physiol. Plant.* 106, 135–141. doi: 10.1034/j.1399-3054.1999.106119.x
- Mohanty, S. P., Hughes, D. P., and Salathé, M. (2016). Using deep learning for image-based plant disease detection. *Front. Plant Sci.* 7:1419. doi: 10.3389/fpls.2016.01419
- Morales, H., Armbrrecht, I., Gonthier, D., and Wyckhuys, K. A. G. (2021). Editorial: moving from a curative to preventative pest management paradigm. *Front. Sustain. Food Syst.* 5:808124. doi: 10.3389/fsufs.2021.808124
- Mustafa, G., Zheng, H., Khan, I. H., Zhu, J., Yang, T., Wang, A., et al. (2024). Enhancing fusarium head blight detection in wheat crops using hyperspectral indices and machine learning classifiers. *Comput. Electron. Agric.* 218:108663. doi: 10.1016/j.compag.2024.108663
- Nagasubramanian, K., Jones, S., Sarkar, S., Singh, A. K., Singh, A., and Ganapathysubramanian, B. (2018). Hyperspectral band selection using genetic algorithm and support vector machines for early identification of charcoal rot disease in soybean stems. *Plant Methods* 14:86. doi: 10.1186/s13007-018-0349-9
- Nagasubramanian, K., Jones, S., Singh, A. K., Sarkar, S., Singh, A., and Ganapathysubramanian, B. (2019). Plant disease identification using explainable 3D deep learning on hyperspectral images. *Plant Methods* 15:98. doi: 10.1186/s13007-019-0479-8
- Namayanandeh, S. M., Kaseb, F., Lesan, S., Rafieian, M., and Khedmat, L. (2013). Olive and sesame oil effect on lipid profile in hypercholesterolemic patients, which better? *Int. J. Prev. Med.* 4, S1059–S1063. doi: 10.4103/2008-7802.120231
- Namiki, M. (2007). Nutraceutical functions of sesame: a review. *Crit. Rev. Food Sci. Nutr.* 47, 651–673. doi: 10.1080/10408390600919114

- Nguyen, C., Sagan, V., Maimaitiyiming, M., Maimaitijiang, M., Bhadra, S., and Kwasniewski, M. T. (2021). Early detection of plant viral disease using hyperspectral imaging and deep learning. *Sensors* 21:742. doi: 10.3390/s21030742
- Oboulbiga, E. B., Douamba, Z., Compaoré-Séréme, D., Semporé, J. N., Dabo, R., Semde, Z., et al. (2023). Physicochemical, potential nutritional, antioxidant and health properties of sesame seed oil: a review. *Front. Nutr.* 10:1127926. doi: 10.3389/fnut.2023.1127926
- Qin, F., Wang, H., Jiang, Q., and Wang, H. (2025). A novel method based on lesion expansion to assess plant disease severity. *Front. Plant Sci.* 16:1510663. doi: 10.3389/fpls.2025.1510663
- Ram, B. G., Oduor, P., Igathinathane, C., Howatt, K., and Sun, X. (2024). A systematic review of hyperspectral imaging in precision agriculture: analysis of its current state and future prospects. *Comput. Electron. Agric.* 222:109037. doi: 10.1016/j.compag.2024.109037
- Rouse, J. W., Haas, R. H., Deering, D. W., Schell, J. A., and Harlan, J. C. (1974). Monitoring the vernal advancement and Retrogradation (green wave effect) of natural vegetation. NASA/GSFC Type III Final Report. Greenbelt, MD, USA: National Aeronautics and Space Administration. Nasa.
- Rumpf, T., Mahlein, A. K., Steiner, U., Oerke, E. C., Dehne, H. W., and Plümer, L. (2010). Early detection and classification of plant diseases with support vector machines based on hyperspectral reflectance. *Comput. Electron. Agric.* 74, 91–99. doi: 10.1016/j.compag.2010.06.009
- Sumabat, L. G., Kemeraite, R. C. Jr., and Brewer, M. T. (2018). Phylogenetic diversity and host specialization of *Corynespora cassiicola* responsible for emerging target spot disease of cotton and other crops in the southeastern United States. *Phytopathology* 108, 892–901. doi: 10.1094/PHYTO-12-17-0407-R
- Szymańska, E., Saccenti, E., Smilde, A. K., and Westerhuis, J. A. (2012). Double-check: validation of diagnostic statistics for PLS-DA models in metabolomics studies. *Metabolomics* 8, 3–16. doi: 10.1007/s11306-011-0330-3
- Wei, P., Zhao, F., Wang, Z., Wang, Q., Chai, X., Hou, G., et al. (2022). Sesame (*Sesamum indicum* L.): a comprehensive review of nutritional value, phytochemical composition, health benefits, development of food, and industrial applications. *Nutrients* 14:4079. doi: 10.3390/nu14194079
- Yoon, H. I., Lee, H., Yang, J. S., Choi, J. H., Jung, D. H., Park, Y. J., et al. (2023). Predicting models for plant metabolites based on PLSR, AdaBoost, XGBoost, and LightGBM algorithms using hyperspectral imaging of *Brassica juncea*. *Agriculture* 13:1477. doi: 10.3390/agriculture13081477
- Zhao, Y. R., Li, X., Yu, K. Q., Cheng, F., and He, Y. (2016). Hyperspectral imaging for determining pigment contents in cucumber leaves in response to angular leaf spot disease. *Sci. Rep.* 6:27790. doi: 10.1038/srep27790
- Zhao, X., Lin, C. W., Wang, J., and Oh, D. H. (2014). Advances in rapid detection methods for foodborne pathogens. *J. Microbiol. Biotechnol.* 24, 297–312. doi: 10.4014/jmb.1310.10013
- Zhou, R. Q., Jin, J. J., Li, Q. M., Su, Z. Z., Yu, X. J., Tang, Y., et al. (2018). Early detection of *Magnaporthe oryzae*-infected barley leaves and lesion visualization based on hyperspectral imaging. *Front. Plant Sci.* 9:1962. doi: 10.3389/fpls.2018.01962
- Zhu, H., Chu, B., Zhang, C., Liu, F., Jiang, L., and He, Y. (2017). Hyperspectral imaging for presymptomatic detection of tobacco disease with successive projections algorithm and machine-learning classifiers. *Sci. Rep.* 7:4125. doi: 10.1038/s41598-017-04501-2