Contents lists available at Science-Gate

International Journal of Advanced and Applied Sciences

Journal homepage: http://www.science-gate.com/IJAAS.html

A garlic disease identification model based on near-infrared spectroscopy and an optimized ResNet

Rongfeng Zhang [1, *], Tang Ye [1], Zexi Li [1], Ting Deng [2]

[1]College of Artificial Intelligence, Zhongkai University of Agriculture and Engineering, Guangzhou, China
[2]Network and Information Technology Office, South China University of Technology, Guangzhou, China

ARTICLE INFO

Article history:
Received 21 May 2025
Received in revised form
28 September 2025
Accepted 14 October 2025

Keywords:
Garlic disease
Near-infrared spectroscopy
Convolutional neural network
Deep learning
Agricultural diagnosis

ABSTRACT

Early detection of garlic diseases is essential for improving agricultural quality and productivity. This study presents a novel garlic disease identification model based on near-infrared (NIR) spectroscopy and a convolutional neural network, named ST-1DResNet (One-dimensional Residual Networks with Squeeze-and-Excitation and tanh activation). The model overcomes the vanishing gradient problem, adaptively adjusts channel weights, and efficiently extracts spectral features without requiring preprocessing or manual feature extraction. Experimental results show that ST-1DResNet achieves a classification accuracy of 97.75%, outperforming the original ResNet and four classical deep learning models by an average of 6.40%. Compared with traditional machine learning methods and optimized SVM models, it improves accuracy by 11.63% and 2.67%, respectively. The model is compact, computationally efficient, and supports fast training, making it suitable for deployment in resource-limited environments. Its strong generalization performance, validated using an external mango dataset, highlights its scalability. Overall, ST-1DResNet provides a practical, accurate, and non-destructive approach for crop disease detection, contributing to quality control and intelligent diagnosis in modern agriculture.

1. Introduction

Garlic is not only a common seasoning that imparts a distinctive flavor to food but also possesses various pharmacological effects, including antibacterial, antiviral, antitumor, and hypoglycemic properties (Tudu et al., 2022). However, during its growth cycle, garlic is highly susceptible to diseases that severely impact both yield and quality, resulting in significant economic losses for farmers (Anum et al., 2024). Garlic cultivation was frequently affected by several pathogenic diseases, including root rot, leaf blight, gray leaf spot, purple spot, and phytophthora (Dedecan et al., 2022). These phytopathogens primarily infected and damaged the foliage, pseudo stems, and bulbs of garlic, ultimately compromising plant growth and significantly reducing crop yield. Specifically, root rot manifested as characteristic symptoms initially appearing at the pseudo stem base and bulb. The infection process began with water-soaked lesions that progressively darkened, developed necrotic ulcerations, and eventually led to tissue maceration (Gálvez and Palmero, 2021). Disease progression typically resulted in complete bulb rot under favorable environmental conditions. Leaf blight primarily infected the foliar tissues and floral stalks of garlic plants. Initial symptoms appeared as small, white necrotic lesions that subsequently expanded into elongated, stripe-like patterns or developed characteristic purple-hued spots (Sharifi et al., 2021). Similarly, purple spot disease predominantly affected leaves and floral stalks. Early-stage infections presented as slightly depressed, whitish lesions with distinct purple pigmentation at the center. As the disease progressed, these lesions enlarged into fusiform (spindle-shaped) necrotic areas with defined margins. Phytophthora primarily affected garlic foliar tissues, with initial infection typically manifesting as chlorotic (yellow-white) water-soaked lesions at either leaf tips or mid-leaf regions (Anum et al., 2024). These symptomatic areas progressively expanded, leading to extensive necrotic foliar damage. The pathological impact of

* Corresponding Author.
Email Address: rongfeng66@zhku.edu.cn (R. Zhang)
https://doi.org/10.21833/ijaas.2025.11.007
Corresponding author's ORCID profile:
https://orcid.org/0000-0003-3711-176X

these garlic diseases was economically significant, causing substantial reductions in both crop yield. Such losses directly compromised grower profitability, particularly in intensive cultivation systems where disease pressure was high (Workneh et al., 2024).

Surveys have shown that disease-related losses typically account for 10%–20% of garlic production, with losses exceeding 30% or even higher in particularly severe years and regions. Garlic leaf blight represented the most destructive foliar disease in garlic cultivation (Zewde et al., 2007). Epidemiological studies demonstrated that severe outbreaks typically resulted in 20-30% yield reduction when control measures were delayed (Cui et al., 2024). In cases of acute infection, the pathogen frequently induced premature plant senescence and garlic scape rot, leading to significant production losses that directly impacted grower profitability. Garlic root rot in mildly affected areas generally reduced yields by 20% to 30%, whereas in regions with approximately 10 years of continuous garlic cultivation, yield losses exceeded 30%. These pathological effects brought serious economic losses to garlic farmers (Shagun et al., 2024). Therefore, the ability to rapidly and accurately identify garlic diseases at an early stage is crucial for implementing timely control measures, minimizing damage, and ensuring the sustainable development of the garlic industry (Vijaykumar et al., 2023).

In the early stages of infection, symptomatic manifestations in garlic plants are often subtle, making it difficult for traditional machine vision systems to achieve effective recognition. Moreover, manual disease identification by non-specialists tends to suffer from low accuracy and is constrained by high labor intensity and low efficiency. Accordingly, it becomes crucial to develop a straightforward, non-damaging, and efficient technique for detecting garlic diseases in their initial stages.

In recent years, NIRS has been widely applied in crop variety identification (Feng et al., 2021). For instance, Sen et al. (2024) employed a characteristic wavelength selection method based on adaptive sliding window permutation entropy (ASW-PE) combined with various classification models to classify rice samples, achieving an accuracy rate of 95%. Similarly, Shao et al. (2017) utilized NIR spectroscopy in combination with Linear Discriminant Analysis (LDA), Quadratic Discriminant Analysis (QDA), and Support Vector Machine (SVM) to classify papaya samples, reaching classification accuracies of 94%, 96%, and 98%, respectively. Zhai et al. (2024) used the Competitive Adaptive Reweighted Sampling (CARS) algorithm to select characteristic wavelengths and applied Least-Squares Support Vector Machine (LS-SVM) modeling to classify rice samples from different growth stages, achieving a correct classification rate of 91.67%.

To enhance the predictive capability of models when dealing with uncertain samples, current research primarily focuses on reducing data dimensions, feature extraction, and data modeling techniques (Liu et al., 2024). While conventional chemometric techniques have been widely applied in NIRS analysis, they continue to encounter significant obstacles and constraints (Zou et al., 2025). These include complex operational procedures, insufficient model robustness, and an inability to fully meet the growing demand for high-accuracy and intelligent early disease detection.

Deep learning technologies, by contrast, eliminate the need for complex data preprocessing and feature extraction, allowing models to be trained directly on raw data (Chen et al., 2024). Deep learning approaches have been successfully applied across a wide range of fields, including natural language processing (Yeboah and Baz Musah, 2022), image processing (Archana and Jeevaraj, 2024), intelligent manufacturing (Neupane et al., 2021), and signal identification (Tian et al., 2022). Researchers have also explored the integration of deep learning techniques with NIRS analysis for rapid plant disease identification. For example, Feng et al. (2024) proposed a model combining deformable convolution and dilated convolution neural networks to detect Asian soybean rust, achieving an overall accuracy of 96.73%. Ong et al. (2025) integrated convolutional neural networks with continuous wavelet transform (CWT) spectrograms and demonstrated that the Random Forest (RF) model incorporating spectral derivative features achieved the best performance in sugarcane disease identification, with an accuracy of 94.87%. Therefore, the combination of convolutional neural networks (CNN) with spectral analysis techniques has proven effective in improving the accuracy and efficiency of plant disease recognition while reducing manual intervention and associated costs.

At present, there are very few studies on the use of NIRS combined with CNN for the identification of garlic diseases. NIR spectral data are rich in chemical and physical information, and compared with two-dimensional image data, NIR datasets are more lightweight and require less computational resources (Shen et al., 2025). Hyperspectral image data demonstrated superior capability for substance characterization compared to conventional two-dimensional RGB images, as the acquired spectral signatures contained significantly more biochemical information. This enhanced data dimensionality enabled more accurate substance identification and classification, since different materials exhibited unique spectral fingerprints across the measured wavelength range. Hence, employing an improved one-dimensional CNN model to classify and detect garlic based on NIR spectral data is both feasible and advantageous.

2. Materials and methods

2.1. Data collection

Garlic samples were collected from the garlic planting base in Da'an Town, Lufeng City, Guangdong

Province, including 12 healthy plants, and plants infected with root rot (5 plants), leaf blight (10 plants), phytophthora (10 plants), gray leaf spot (8 plants), and purple spot (9 plants), totaling 54 garlic samples. Fig. 1 shows the RGB images of various early-stage garlic diseases taken at the planting base.

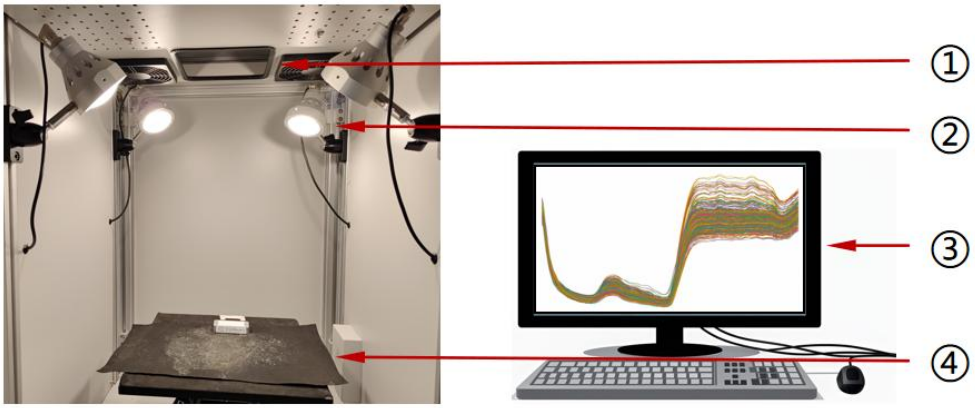
Fig. 1: Images of different garlic diseases

To collect the infrared spectral data of garlic, we employed a hyperspectral image acquisition system consisting of a near-infrared hyperspectral camera (GaiaSky mini, Dualix, Sichuan, China) and a specially designed dark chamber, as illustrated in Fig. 2. The hyperspectral camera was fixed at the top of the dark chamber, covering a spectral range from 386.7 nm to 1,016.7 nm, with an image resolution of 696 × 700 pixels. Inside the dark chamber, a customized scanning platform and a lighting system composed of four 50W tungsten-halogen lamps were installed, with the light sources positioned at a 45-degree angle relative to the sample. During the spectral acquisition process, garlic samples were placed flat on the lifting platform, and data were collected using diffuse reflectance. The exposure time was set to 9.98 ms, and the scanning speed was set to 0.8 mm/s. Fig. 3 presents the hyperspectral images of garlic samples affected by different diseases.

A total of 591 near-infrared spectral data points were collected from different positions of leaves from the 54 garlic samples. The original spectral data are shown in Fig. 4. Table 1 summarizes the distribution of sample numbers and spectral data across different garlic disease categories.

Table 1: Distribution of sample numbers and spectral data by garlic disease category

Garlic category	Number of samples	Number of spectral variables
Healthy	97	256
Root rot	84	256
Gray leaf spot	102	256
Leaf blight	102	256
Phytophthora	102	256
Purple spot	104	256



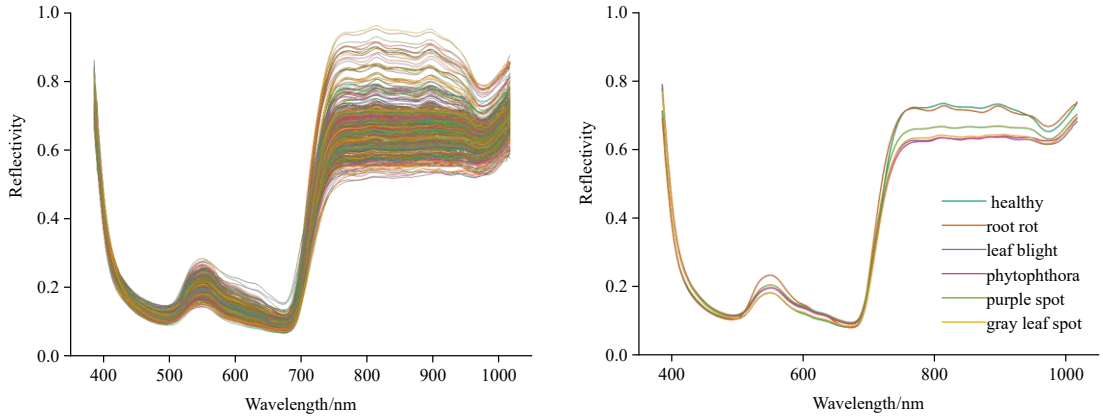
①: Hyperspectral camera; ②: Light source; ③: Computer; ④: Scanning platform

Fig. 2: Hyperspectral image acquisition system

Fig. 3: Hyperspectral images of different garlic diseases

2.2. Garlic disease identification model based on improved ResNet

The objective of this research is to accomplish swift and precise detection of garlic diseases through the establishment of six deep learning models: VGG, ResNet, Xception, MobileNet, DenseNet, and a model named ST-1DResNet, which incorporates a channel attention mechanism. The structure of the ST-1DResNet model is shown in Fig. 5.



(a) Raw spectra of garlic samples (b) Average spectra of garlic samples

Fig. 4: Raw NIR spectra of different garlic disease categories

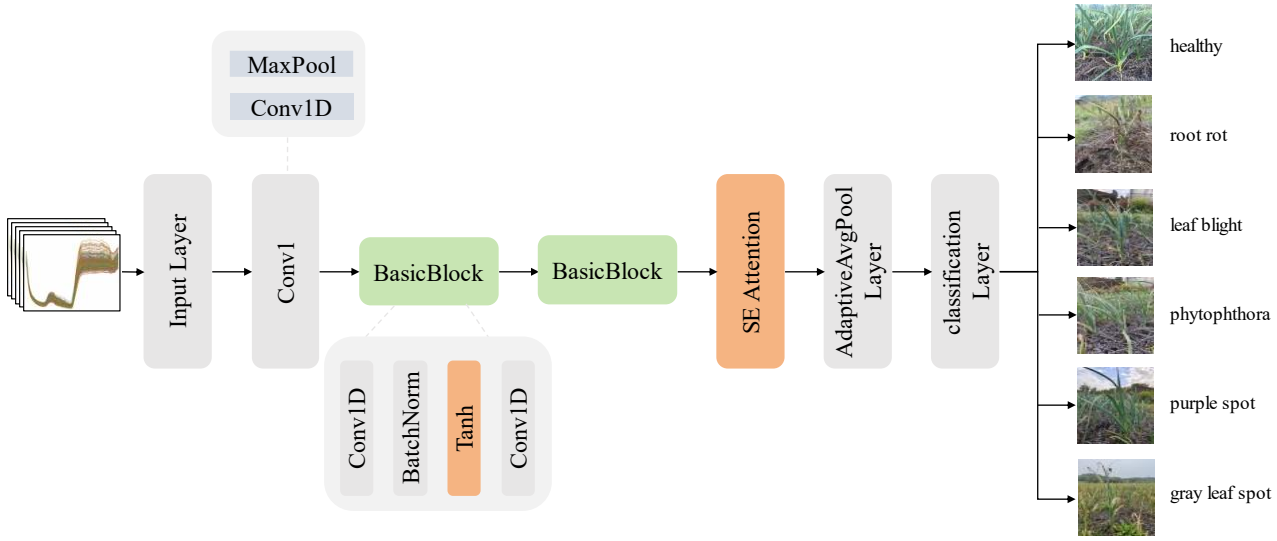


Fig. 5: Structure of the ST-1DResNet model

2.2.1. Deep learning models

CNN is a common type of deep learning model, mainly employed for tasks in computer vision like object recognition, image recognition, and classification (Zhang et al., 2024). The essence of CNN lies in utilizing convolutional layers to extract features, employing pooling layers to decrease data dimensions and condense information, and ultimately accomplishing image classification or regression via fully connected layers (Muhammad et al., 2022).

One-dimensional convolution, or 1D convolution, primarily handles one-dimensional data streams or sequences, with both the input and the convolutional kernel being one-dimensional. By sliding the convolution kernel along the input data and performing element-wise multiplication and summation operations, 1D convolution computes a single output value, thus capturing local features in the sequential data (Ma et al., 2023).

In contrast, two-dimensional convolution is mostly utilized for handling two-dimensional data like images. In this case, the convolution kernel is a two-dimensional matrix that slides over each local region of the input data, performing element-wise multiplication and summation operations to generate a single output value (Promboonruang and Boonrod, 2023). 2D convolution effectively captures spatial features in images, such as edges and textures.

The VGG network, proposed by the Visual Geometry Group at the University of Oxford in 2014 (Simonyan and Zisserman, 2014), is a deep convolutional neural network architecture characterized by the use of small-sized convolution kernels and a deeper network structure. This design significantly enhances the model's ability to extract image features and improves its capability to fit complex patterns by increasing network depth.

ResNet, proposed by He et al. (2016) at Microsoft Research in 2015, is a deep convolutional neural network architecture. In traditional CNN, as network depth increases, the training process often becomes complex and difficult to optimize due to vanishing or exploding gradients. However, ResNet effectively addresses this problem by introducing skip connections (also called shortcut connections) within residual blocks. This design allows gradients to bypass multiple layers, significantly improving the training efficiency and stability of deep networks, thereby enabling more efficient learning and optimization. Xception, inspired by the Inception architecture, is a deep convolutional neural network

proposed by Google in 2017, Deep Learning with Depthwise Separable Convolutions (Chollet, 2017). Its core innovation lies in replacing Inception modules with depthwise separable convolutions, thereby improving model performance without significantly increasing computational complexity.

MobileNet, proposed by Google in 2017, is a lightweight convolutional neural network architecture (Howard et al., 2017) designed for mobile and embedded devices. Its key innovation is the introduction of depthwise separable convolutions, which break down traditional convolution operations into two steps: Depthwise convolution and pointwise convolution. This design significantly reduces the number of parameters and computational complexity while maintaining high accuracy.

DenseNet, introduced by the Google Brain team in 2016 (Huang et al., 2017), is a deep learning architecture aimed at addressing the information flow and vanishing gradient problem in traditional CNN. DenseNet's core idea is dense connectivity, where each layer is connected to every other preceding layer. This approach maximizes the use of features extracted by earlier layers, enhances feature propagation, mitigates vanishing gradient problems, and promotes feature reuse.

2.2.2. Squeeze-and-excitation module

In deep learning, CNN typically extract features through convolution operations, but this process often mixes spatial and channel information, leading to insufficient modeling of inter-channel relationships. To address this, the Squeeze-and-Excitation (SE) module was proposed (Hu et al., 2018), aiming to explicitly model the interdependencies among feature channels.

The core idea of the SE module is to dynamically recalibrate channel-wise feature responses by learning the importance of each channel. Specifically, the SE module comprises three main steps: Squeeze, Excitation, and Scale.

- Squeeze: Global average pooling compresses the spatial information of each channel into a single scalar, capturing global context.
- Excitation: Two fully connected layers and non-linear activation functions (ReLU and Sigmoid) are used to generate channel-wise weights that reflect each channel's importance.
- Scale: The generated weights are applied to the original feature maps to emphasize important features and suppress less useful ones.

Essentially, the SE module implements an attention mechanism along the channel dimension, enabling the model to focus more effectively on significant feature channels. The structure of the SE attention mechanism is illustrated in Fig. 6.

2.2.3. Tanh activation function

Tanh (Hyperbolic Tangent Function) is a commonly used activation function (Li and Ni, 2024), with a domain of R and an output range of (-1, 1). Its output is centered around zero, which helps accelerate neural network training by ensuring more balanced gradient updates, thereby alleviating the vanishing gradient problem. As a non-linear function, Tanh enables neural networks to learn complex features and patterns. Its mathematical expression is shown in Eq. 1:

$$Tanh\,(x) = \frac{e^{x}-e^{-x}}{e^{x}+e^{-x}} \quad (1)$$

2.3. Model evaluation metrics

In machine learning and deep learning, model recognition accuracy is typically evaluated using metrics such as accuracy, precision, recall, F1 score, and the kappa coefficient. By analyzing changes in these metrics, corresponding model optimization strategies can be adjusted. These standards help assess model performance, identify problems, and facilitate performance improvement.

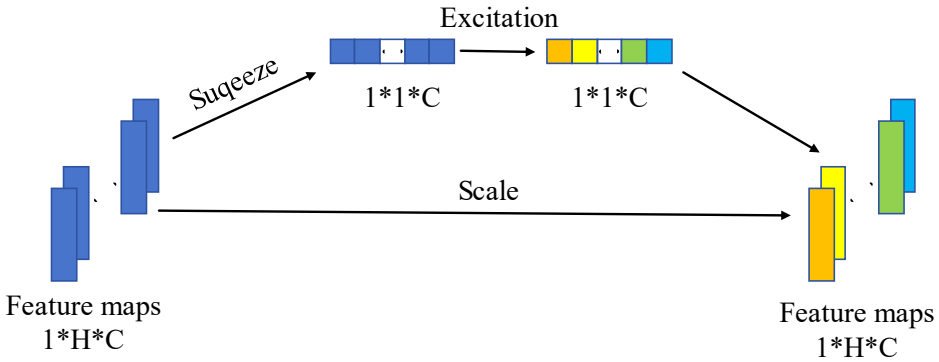


Fig. 6: Structure of the SE attention mechanism

Accuracy (Ninh et al., 2024) is a metric used to measure model performance, which is calculated by dividing the number of correct predictions by the total number of samples. The calculation equation for Accuracy is as follows:

$$\text{Accuracy} = \frac{\text{TP+TN}}{\text{TP+TN+FP+FN}} \quad (2)$$

Precision (Liu et al., 2025) measures the proportion of true positive predictions among all

positive predictions, reflecting the model's accuracy in positive predictions.

Recall (Chouhan et al., 2024) refers to the proportion of all actual positives that were classified correctly as positives, reflecting the model's ability to identify positive samples.

F1 score (Chouhan et al., 2024) is the harmonic mean of Precision and Recall, providing a balance between the two metrics.

The kappa coefficient (Martín Andrés and Álvarez Hernández, 2025) is used to evaluate the performance of a classification model by considering the effect of random chance, reflecting the difference between the model's classification results and results expected by random chance.

The calculation formulas for Precision, Recall, F1 score, and kappa are shown in Eqs. 3, 4, 5, and 6:

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

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

$$\text{F1 score} = \frac{2\times\text{recall}\times\text{precision}}{\text{recall+precision}} \quad (5)$$

$$\text{Kappa} = \frac{\text{p}_\text{o}-\text{p}_\text{e}}{1-\text{p}_\text{e}} \quad (6)$$

In the aforementioned formulas:

- TP (True Positive) refers to the number of actual positive samples correctly predicted as positive;
- FN (False Negative) refers to the number of actual positive samples incorrectly predicted as negative;
- TN (True Negative) refers to the number of actual negative samples correctly predicted as negative;
- FP (False Positive) refers to the number of actual negative samples incorrectly predicted as positive;

Here, p_o corresponds to the proportion of observed agreement, while p_e indicates the expected chance of agreement.

3. Results

All experimental results presented below were obtained under the Windows 11 operating system, using a CPU of i5-12600KF and a GPU of NVIDIA GeForce 4060Ti 16G. The code for constructing and optimizing all models was run in the PyTorch 1.13.1 environment, with Python version 3.8.

3.1. Analysis of machine learning optimization models

In this section, three optimization algorithms were implemented: GA-SVM, PSO-SVM, and CFOA-SVM. To accurately compare the classification performance of each optimization algorithm, widely-applied machine learning classification models were chosen for performance comparison. These models include Support Vector Machine (SVM), Random Forest (RF), K-Nearest Neighbors (KNN), Decision Tree (DT), and Multilayer Perceptron (MLP). The classification performance was evaluated using metrics such as accuracy, precision, recall, F1 score, kappa, and the confusion matrix. The confusion matrix was primarily used to compare the objective results against the actual categories. The validation methods all use ten-fold cross-validation. The performance metrics and confusion matrices for SVM, RF, KNN, DT, GA-SVM, PSO-SVM, and CFOA-SVM were calculated and are presented in Table 2 and Fig. 7.

From the result statistics in Table 2, it could be observed that during the establishment of the SVM model for garlic disease classification, the CFOA-SVM model achieved the highest classification accuracy of 95.63%. Under identical test samples and search conditions, the CFOA algorithm demonstrated faster and more optimized performance in identifying SVM's two key parameters. Consequently, the CFOA-SVM model significantly outperformed both GA-SVM and PSO-SVM in garlic disease identification. According to the confusion matrix results, the CFOA-SVM model achieved a 100% correct recognition rate for root rot and purple spot, while the accuracy for all other categories exceeded 87%. Specifically, 3.23% of gray leaf spot samples were misclassified as leaf blight, 9.68% and 3.23% of leaf blight samples were misclassified as Phytophthora and purple spot, respectively, 3.12% of purple spot samples were misclassified as leaf blight, and 3.85% of root rot samples were not correctly identified. The primary reason for this misclassification lies in the highly similar near-infrared spectral data, where differences in the chemical composition and physical characteristics of garlic samples are not sufficiently distinct.

Table 2: Comparison of prediction results for machine learning models

Methods	Accuracy (%)	Precision (%)	Recall (%)	F1 score (%)	Kappa (%)
SVM	86.41	84.34	82.80	82.49	78.86
RF	89.51	90.44	89.79	89.63	87.39
KNN	87.84	88.88	88.22	87.95	85.39
DT	71.90	73.42	72.59	72.34	66.26
MLP	94.96	95.00	94.96	94.95	93.94
GA-SVM	95.03	95.15	95.04	95.03	94.03
PSO-SVM	94.59	94.76	94.70	94.70	93.51
CFOA-SVM	95.63	95.75	95.63	95.63	94.53

To further improve the precision and stability of garlic disease classification, we proposed a deep learning model, ST-1DResNet, which leverages the powerful feature extraction capabilities of CNN to capture subtler features within the near-infrared spectral data of garlic samples.

3.2. Analysis of deep learning optimization models

To avoid potential category imbalance problems, the synthetic minority oversampling technique (SMOTE) method was employed to balance the

category distributions before training, preventing model bias from imbalanced classes. This research utilized the widely adopted categorical cross-entropy (CCE) loss function to assess the discrepancy between predicted and actual outcomes. The formula for the cross-entropy loss function is given in Eq. 7:

$$\text{Loss} = \frac{1}{N}\sum_{I} L_i = -\frac{1}{N}\sum_{i}^{M}\sum_{c=1}^{M} y_{ic}\log(p_{ic}) \quad (7)$$

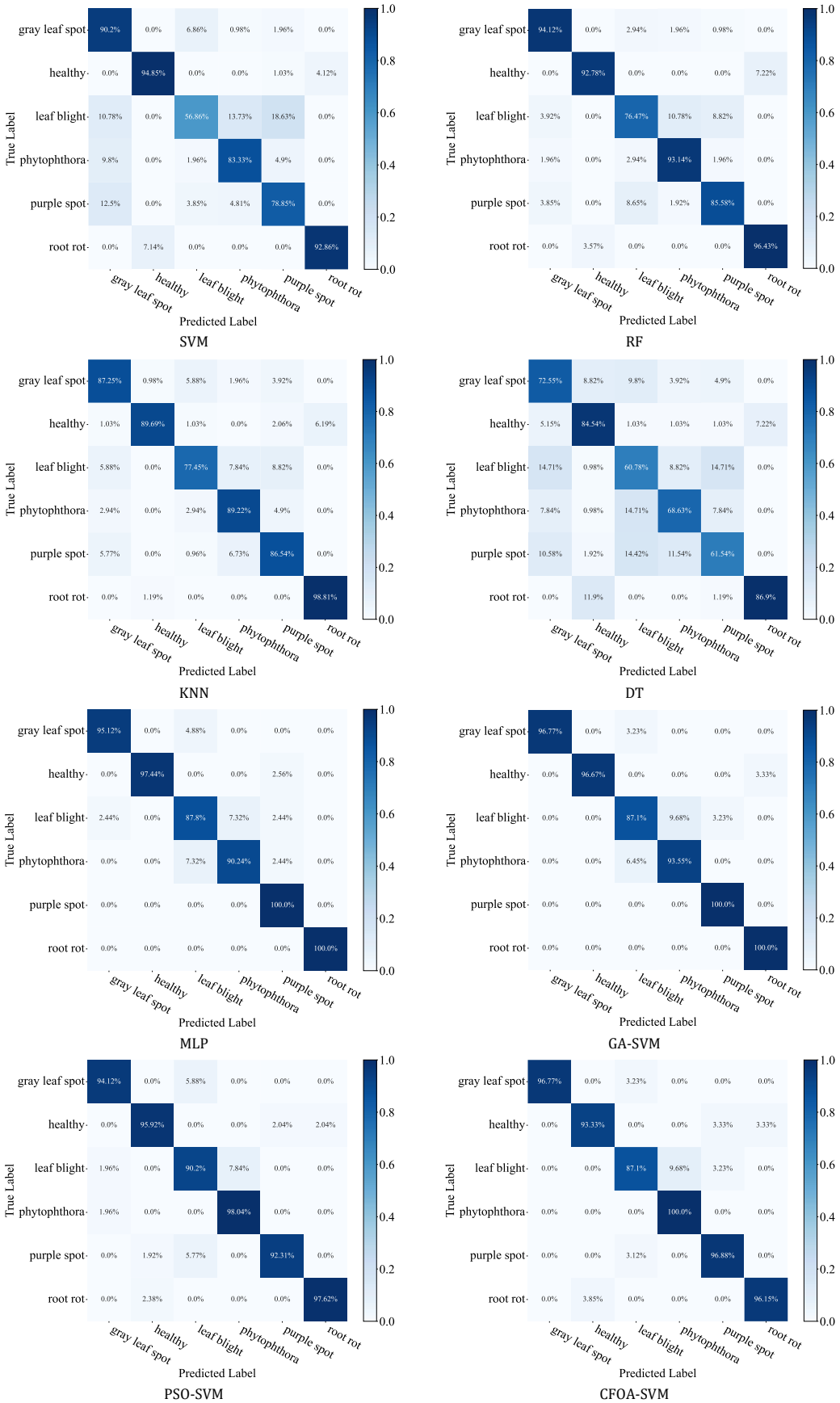


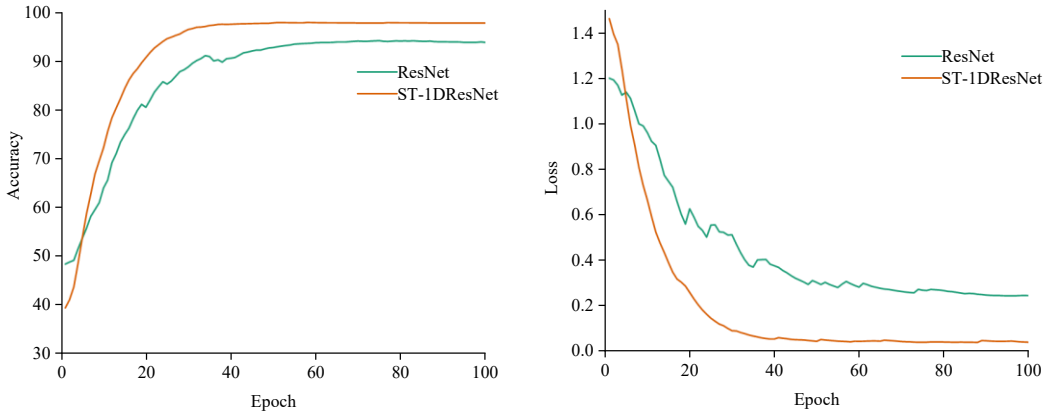
Fig. 7: Confusion matrices for machine learning models

During training, the Adam optimizer is utilized for model optimization, with a learning rate of 0.001, a batch size of 16, and 100 iterations. Fig. 8 illustrates the changes in loss and accuracy for both the ResNet model and its enhanced ST-1DResNet model on the validation set throughout the training process. With the increase in iterations, the loss of the ST-1DResNet model gradually decreases. Notably, the loss exhibits a rapid decline within the first 40 iterations, after which it stabilizes and approaches zero. Compared to the ResNet model, the ST-1DResNet model demonstrates lower loss, reduced volatility, and more stable convergence.

To assess the performance of the garlic disease identification model on the test set, the results of the original and enhanced models were compared during testing. Fig. 9 presents the confusion matrices of the models, which assess classification performance and illustrate the relationship between predicted and actual results for each category in matrix form. In these matrices, the blue diagonal represents correctly classified results, while all other cells indicate misclassifications. The deeper the color, the better the classification performance of the model. In the ResNet model, the classification of root rot was completely accurate, but the other five categories exhibited varying degrees of confusion. Specifically, 3.7% of healthy samples were misclassified as root rot; 10.34% of leaf blight samples were misclassified as phytophthora; and 6.06%, 3.03%, and 6.06% of phytophthora samples were misclassified as leaf blight, purple spot, and gray leaf spot, respectively. Additionally, 2.7% of purple spot samples were misclassified as gray leaf spot, and 5.88% of gray leaf spot samples were misclassified as phytophthora.

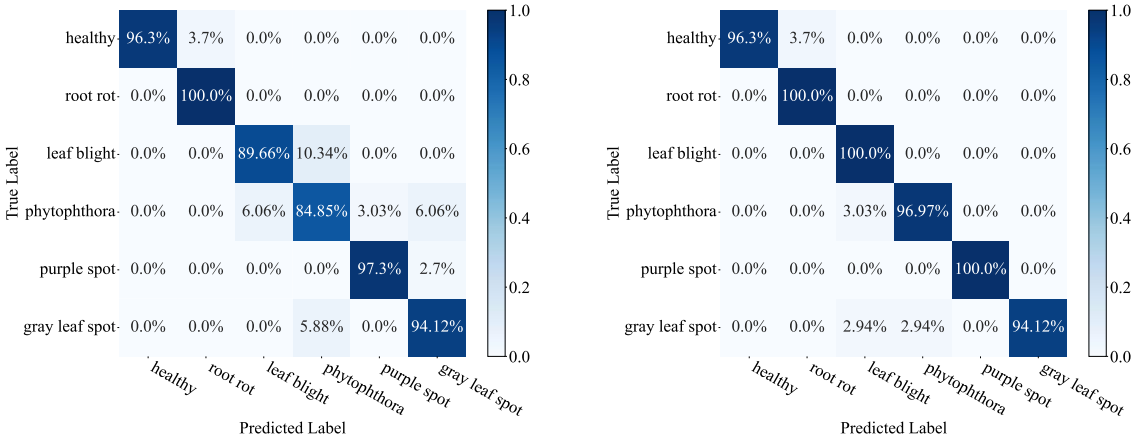
In contrast, in the ST-1DResNet model, the classifications of root rot, leaf blight, and purple spot achieved 100% accuracy. However, 3.7% of healthy samples were still misclassified as root rot, and 3.03% of Phytophthora samples were misclassified as leaf blight. Furthermore, 2.94% and another 2.94% of gray leaf spot samples were misclassified as leaf blight and phytophthora, respectively. The primary reason for these misclassifications is that leaf blight, phytophthora, and gray leaf spot share highly similar spectral characteristics, making them harder to distinguish.

Compared to the original model, the ST-1DResNet model captures the spectral features of garlic diseases in greater detail, thus improving classification accuracy.



(a) Accuracy curve (b) Loss curve

Fig. 8: Accuracy and loss curves of ResNet and ST-1DResNet



(a) ResNet (b) ST-1DResNet

Fig. 9: Comparison of confusion matrices between ResNet and ST-1DResNet

4. Discussion

4.1. Analysis of model testing results

To further evaluate the performance of the proposed ST-1DResNet model, a comparison was made with the aforementioned machine learning models and their optimized versions. The comparison results are shown in Table 3. As can be seen from Table 3, the ST-1DResNet model achieved an accuracy of 97.75%, which outperformed all other machine learning models. In addition to accuracy, ST-1DResNet also achieved the highest values in precision, recall, f1 score, and kappa coefficient among all models.

These experimental results demonstrate that the ST-1DResNet model not only delivers higher prediction accuracy for garlic disease identification compared to traditional machine learning models but also exhibits superior robustness and reliability.

To test the performance of the attention mechanism and activation function within our proposed ST-1DResNet model, we performed comparative studies against various other deep learning models. The models used for comparison include VGG, ResNet, Xception, MobileNet, DenseNet, and ST-1DResNet. The classification results are shown in Table 4. Table 4 presents the classification prediction results and model complexity evaluations for various deep learning models. It can be seen that the attention mechanism-based ST-1DResNet model achieves an accuracy of 97.75%, which markedly surpasses the outcomes achieved with conventional machine learning and deep learning approaches discussed previously. This improvement is attributed to the Tanh activation function used in the residual connections, which enhances the model's stability and progressively amplifies feature effects during training. Additionally, the channel attention module concentrates on significant feature details, thereby improving the model's prediction accuracy. Compared to other deep learning models, the ST-1DResNet model has fewer parameters, with only 3.88M parameters, a model size of 14.79M, and 38.27M FLOPs. Although its FLOPs increase by just 6.77M compared to MobileNet, its accuracy improves by 5.05%. Moreover, the ST-1DResNet model's training time is significantly shorter than that of other deep learning models, requiring only 68.71 seconds. Its testing time is just 0.08 seconds, which is lower than that of other deep learning models.

Table 3: Comparison of prediction results between ST-1DResNet and machine learning models

Methods	Accuracy (%)	Precision (%)	Recall (%)	F1 score (%)	Kappa (%)
SVM	86.41	84.34	82.80	82.49	78.86
RF	89.51	90.44	89.79	89.63	87.39
KNN	87.84	88.88	88.22	87.95	85.39
DT	71.90	73.42	72.59	72.34	66.26
MLP	94.96	95.00	94.96	94.95	93.94
GA-SVM	95.03	95.15	95.04	95.03	94.03
PSO-SVM	94.59	94.76	94.70	94.70	93.51
CFOA-SVM	95.63	95.75	95.63	95.63	94.53
ST-1DResNet	97.75	97.85	97.75	97.76	97.28

Table 4: Comparison of deep learning models' prediction results

Methods	Accuracy (%)	Precision (%)	Recall (%)	F1 score (%)	Kappa (%)	Parameters (M)	Size (M)	FLOPs (M)	Train time (s)	Test time (s)
DenseNet	85.39	85.69	85.39	85.44	82.34	7.44	28.37	112.5	423.27	0.31
VGG	93.26	93.58	93.26	93.30	91.84	11.93	45.54	176.84	79.00	0.09
Xception	92.13	92.77	92.13	92.26	90.50	21.10	80.50	142.21	119.16	0.09
MobileNet	92.70	92.72	92.70	92.64	91.16	8.68	33.10	31.50	179.91	0.12
ResNet	93.26	93.30	93.26	93.26	91.84	3.85	14.67	38.35	101.99	0.08
ST-1DResNet	97.75	97.85	97.75	97.76	97.28	3.88	14.79	38.27	68.71	0.08

To verify the effectiveness of the ST-1DResNet model we constructed and to assess the contribution and impact of each module on the overall performance, we conducted ablation experiments. The results of these experiments are shown in Table 5, where "SE" and "Tanh" refer to the SE attention mechanism and Tanh activation function, respectively.

According to the analysis in Table 5, the ST-1DResNet model achieves the highest classification accuracy. The ablation experiment results indicate that, with a slight increase in model size, both modules contribute to improving the model's generalization performance. Adding only the SE attention mechanism increases accuracy, precision, recall, f1 score, and kappa by 1.68%, 1.77%, 1.68%, 1.67%, and 2.04%, respectively, while reducing training time by 9.3 seconds. After adding only the Tanh activation function, accuracy, precision, recall, f1 score, and kappa increase by 3.37%, 3.41%, 3.37%, 3.37%, and 4.08%, respectively. Meanwhile, the parameters, size, and FLOPs decrease by 0.01M, 0.01M, and 0.02M, respectively, and the training time is reduced by 32.22 seconds.

Compared to the original ResNet model, the combination of both modules in the ST-1DResNet model leads to a slight increase in model size, with parameters and size increasing by 0.03M and 0.12M, respectively. The computational cost slightly decreases, with FLOPs reduced by 0.08M. In terms of time, the training time decreases by 33.28 seconds. However, the model achieves higher accuracy, with accuracy, precision, recall, f1 score, and kappa improving by 4.49%, 4.55%, 4.49%, 4.50%, and 5.44%, respectively. In terms of testing time, testing the 178 sets of data took only 0.08 seconds,

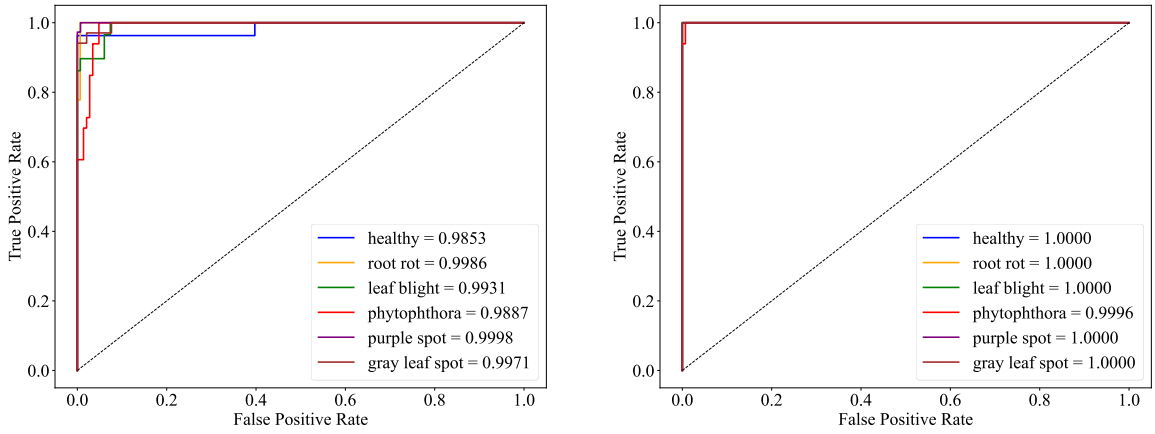
averaging 0.45 milliseconds per set, which is faster than other deep learning models.

After optimizing the model, accuracy was improved without an increase in testing time, indicating that the ST-1DResNet model is suitable for application scenarios requiring rapid and accurate identification. Therefore, by integrating the attention mechanism and Tanh activation function into the ResNet model, the ST-1DResNet model achieves higher accuracy in garlic disease identification.

In the ablation experiments, the AUC-ROC curves before and after the model improvement are shown in Fig. 10. In the ST-1DResNet model, the AUC score for phytophthora is 0.9996, while the AUC scores for healthy samples, root rot, leaf blight, purple spot, and gray leaf spot all reach 1.0, indicating the excellent performance of the model.

Table 5: Ablation experiment results

SE	Tanh	Accuracy (%)	Precision (%)	Recall (%)	F1 score (%)	Kappa (%)	Parameters (M)	Size (M)	FLOPs (M)	Train time (s)	Test time (s)
		93.26	93.30	93.26	93.26	91.84	3.85	14.67	38.35	101.99	0.08
√		94.94	95.07	94.94	94.93	93.88	3.88	14.80	38.39	92.69	0.08
	√	96.63	96.71	96.63	96.63	95.92	3.84	14.66	38.23	69.77	0.08
√	√	97.75	97.85	97.75	97.76	97.28	3.88	14.79	38.27	68.71	0.08



(a) ResNet (b) ST-1DResNet

Fig. 10: AUC-ROC curves for ResNet and ST-1DResNet

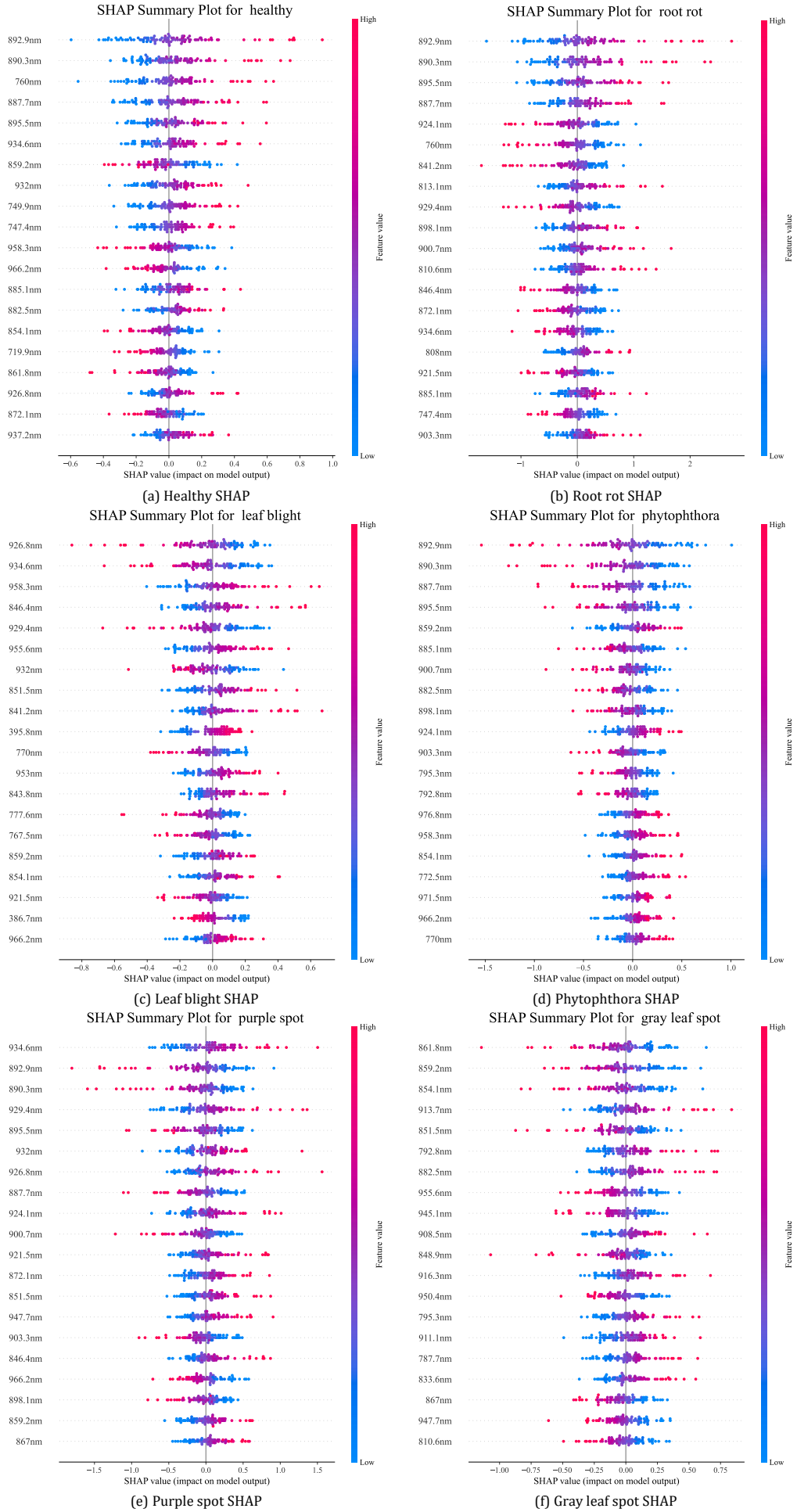
4.2. SHAP interpretability for visualizing feature contributions

SHAP (SHapley Additive exPlanations) values serve as a powerful model interpretation tool, providing new perspectives for understanding and trusting machine learning models (Lundberg and Lee, 2017). The SHAP value is a cooperative game theory-based method for interpreting feature contributions to model outputs. By revealing the decision logic within models, it enhances their transparency and credibility. In this study, we employed the SHAP method to analyze and visualize spectral regions critical for model predictions. The SHAP summary plot (global importance), feature importance ranking, and waterfall plot were generated using the top 20 most significant features selected from 256 spectral bands, ordered by SHAP value importance. These results highlighted the key spectral regions prioritized by the model, as illustrated in Fig. 11 and Fig. 12.

With Fig. 11, it could be observed that the features located at the top of the graph exerted a greater influence on the model predictions, indicating that these features played a key role in predicting garlic disease categories. Taking the 892.9 nm band in Fig. 11a as an example, the red dots were predominantly distributed in the region of positive SHAP values, which suggested that higher reflectance values at 892.9 nm tended to classify garlic samples as healthy, while lower values tended to associate them with diseased categories. Among the remaining 19 feature bands, blue (low values) and red (high values) regions differentially affected positive and negative SHAP values, implying that these features also significantly contributed to the classification behavior. Fig. 11 revealed that 36 features primarily influenced the model predictions, with seven co-occurring key features: 892.9 nm, 890.3 nm, 895.5 nm, 859.2 nm, 924.1 nm, 903.3 nm, and 854.1 nm. Fig. 12 displayed the top 20 important bands ranked by SHAP value importance. SHAP values served as a powerful interpretative tool, quantifying feature contributions to enhance the transparency of the ST-1DResNet model predictions.

4.3. Generalization experiment

To verify the generalization ability of the model proposed in this paper, we used the mango dataset for validation. This dataset comes from the paper by Anderson et al. (2020). The dataset includes ten different mango varieties, such as Calypso, HG, and Keitt, with a total of 12,011 data points. The original spectral data of the mango dataset are shown in Fig. 13. The validation experiment results are shown in Table 6.



(a) Healthy SHAP

(b) Root rot SHAP

(c) Leaf blight SHAP

(d) Phytophthora SHAP

(e) Purple spot SHAP

(f) Gray leaf spot SHAP

Fig. 11: Summary plot of SHAP for the ST-1DResNet model for individual categories

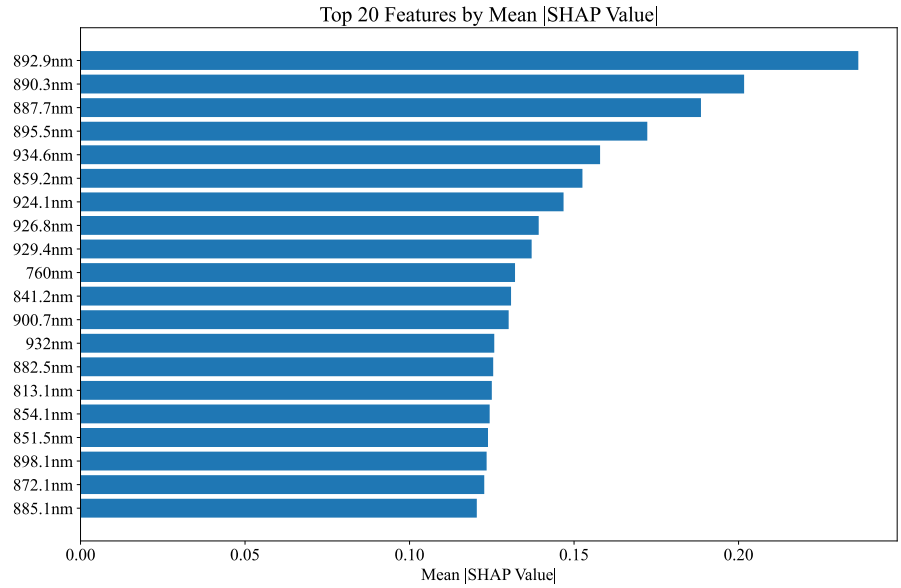
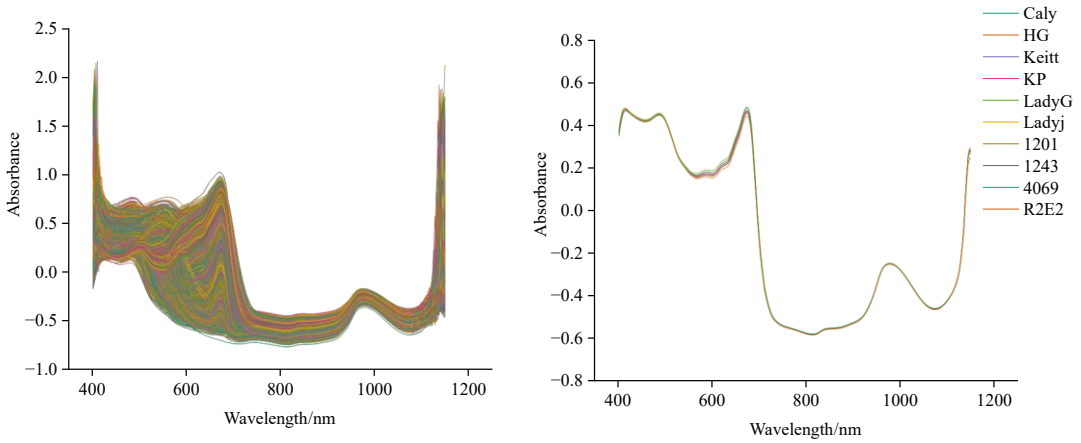


Fig. 12: Importance plot of SHAP features for the ST-1DResNet model



(a) Original mango spectrum (b) Average mango spectrum

Fig. 13: Original NIR spectra of mango varieties

Table 6: Mango dataset validation experiment results

SE	Tanh	Accuracy (%)	Precision (%)	Recall (%)	F1 score (%)	Kappa (%)
		93.73	94.51	93.73	93.63	92.34
√		94.03	94.23	94.03	93.68	92.69
	√	95.17	95.60	95.17	95.25	94.11
√	√	97.84	97.87	97.84	97.83	97.35

As shown in Table 6, the accuracy of the ST-1DResNet model is 97.84%, which is a 4.11% improvement over the ResNet model before optimization, thus validating the generalization ability of the model.

In summary, the ST-1DResNet model, based on the attention mechanism and infrared spectroscopy technology, yields better experimental results. The model achieves an accuracy of 97.75%, precision of 97.85%, recall of 97.75%, f1 score of 97.76%, and kappa coefficient of 97.28%, outperforming the results obtained by traditional machine learning and deep learning models. In contrast to traditional deep learning models, the model constructed in this paper allows the attention mechanism to focus on relevant feature information while suppressing unrelated features by calculating attention parameters. This enhances the extraction of spectral features from garlic, while the Tanh function improves the efficiency of weight updates, solving the vanishing gradient problem, and achieving more accurate experimental results in prediction and classification.

In conclusion, the model proposed in this paper demonstrates better performance in feature extraction and application, leading to significant improvements across all performance evaluation metrics. This model can be applied in fields such as garlic disease identification and variety classification, providing an effective reference value for research in these areas.

5. Conclusion

This paper proposes a garlic disease identification model based on the attention mechanism and NIRS. The combination of NIRS and deep learning methods provides a rapid and efficient approach for the non-destructive detection of garlic diseases, demonstrating its feasibility for the quick identification of garlic diseases. To validate the

effectiveness of the algorithm, we used garlic samples affected by different diseases as the research objects and conducted comparative experiments with several classic deep learning models.

The experimental results show that the model proposed in this paper outperforms the others in terms of performance. This method provides a new, effective, accurate, and non-destructive approach for the identification of plant diseases in agriculture. The model demonstrates good generalization ability and is applicable in many fields. It is not only suitable for garlic disease recognition but also applicable to disease detection in other crops. This contributes to improving the quality control and production efficiency of agricultural products and has a positive impact on the promotion of rapid, non-destructive disease detection technology for crops.

List of abbreviations

1D	One-dimensional
1DResNet	One-dimensional residual networks
AUC	Area under the curve
AUC-ROC	Area under the receiver operating characteristic curve
CARS	Competitive adaptive reweighted sampling
CCE	Categorical cross-entropy
CFOA	Chaotic fruit fly optimization algorithm
CNN	Convolutional neural network
CPU	Central processing unit
CWT	Continuous wavelet transform
DT	Decision tree
FLOPs	Floating point operations
FN	False negative
FP	False positive
GA-SVM	Genetic algorithm-support vector machine
GPU	Graphics processing unit
KNN	K-nearest neighbors
LDA	Linear discriminant analysis
LS-SVM	Least-squares support vector machine
MLP	Multilayer perceptron
NIR	Near-infrared
NIRS	Near-infrared spectroscopy
PSO-SVM	Particle swarm optimization-support vector machine
QDA	Quadratic discriminant analysis
RF	Random forest
RGB	Red, green, blue
SE	Squeeze-and-excitation
SHAP	Shapley additive explanations
SMOTE	Synthetic minority oversampling technique
ST-1DResNet	One-dimensional residual networks with squeeze-and-excitation and tanh activation
SVM	Support vector machine
TN	True negative
TP	True positive
VGG	Visual geometry group

Funding

This work was supported in part by the Science and Technology Commissioner Project of Guangdong Provincial Department of Science and Technology under Grant KTP20240366 and KTP20210242.

Compliance with ethical standards

Conflict of interest

The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

References

Anderson NT, Walsh KB, Subedi PP, and Hayes CH (2020). Achieving robustness across season, location and cultivar for a NIRS model for intact mango fruit dry matter content. Postharvest Biology and Technology, 168: 111202. https://doi.org/10.1016/j.postharvbio.2020.111202

Anum H, Tong Y, and Cheng R (2024). Different preharvest diseases in garlic and their eco-friendly management strategies. Plants, 13(2): 267. https://doi.org/10.3390/plants13020267 **PMid:38256820 PMCid:PMC10818302**

Archana R and Jeevaraj PE (2024). Deep learning models for digital image processing: A review. Artificial Intelligence Review, 57: 11. https://doi.org/10.1007/s10462-023-10631-z

Chen F, Li S, Han J, Ren F, and Yang Z (2024). Review of lightweight deep convolutional neural networks. Archives of Computational Methods in Engineering, 31(4): 1915-1937. https://doi.org/10.1007/s11831-023-10032-z

Chollet F (2017). Xception: Deep learning with depthwise separable convolutions. In the Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, IEEE, Honolulu, USA: 1251-1258. https://doi.org/10.1109/CVPR.2017.195

Chouhan SS, Singh UP, Sharma U, and Jain S (2024). Classification of different plant species using deep learning and machine learning algorithms. Wireless Personal Communications, 136(4): 2275-2298. https://doi.org/10.1007/s11277-024-11374-y

Cui Y, Zhang S, Dong X, Li Q, Huang Y, Meng X, Wang Y, and Ye Y (2024). Garlic-specific fertilizer improves economic and environmental outcomes in China. HortScience, 59(5): 605-612. https://doi.org/10.21273/HORTSCI17677-23

Dedecan O, Talapov T, Demral M, Sarpkaya K, Ceyhan Dİ, and Can C (2022). Molecular and pathogenic characterization of *Fusarium oxysporum* and *Fusarium proliferatum* causing basal root rot in garlic in Turkey. Australasian Plant Disease Notes, 17: 37. https://doi.org/10.1007/s13314-022-00484-w

Feng J, Zhang S, Zhai Z, Yu H, and Xu H (2024). DC2Net: An Asian soybean rust detection model based on hyperspectral imaging and deep learning. Plant Phenomics, 6: 0163. https://doi.org/10.34133/plantphenomics.0163 **PMid:38586218 PMCid:PMC10997487**

Feng L, Wu B, Zhu S, He Y, and Zhang C (2021). Application of visible/infrared spectroscopy and hyperspectral imaging with machine learning techniques for identifying food varieties and geographical origins. Frontiers in Nutrition, 8: 680357. https://doi.org/10.3389/fnut.2021.680357 **PMid:34222304 PMCid:PMC8247466**

Gálvez L and Palmero D (2021). Incidence and etiology of postharvest fungal diseases associated with bulb rot in garlic (*Alllium sativum*) in Spain. Foods, 10(5): 1063. https://doi.org/10.3390/foods10051063 **PMid:34065850 PMCid:PMC8151520**

He K, Zhang X, Ren S, and Sun J (2016). Deep residual learning for image recognition. In the Proceedings of the IEEE Conference

on Computer Vision and Pattern Recognition, IEEE, Las Vegas, USA: 770-778. https://doi.org/10.1109/CVPR.2016.90 **PMid:26180094**

Howard AG, Zhu M, Chen B, Kalenichenko D, Wang W, Weyand T, Andreetto M, and Adam H (2017). Mobilenets: Efficient convolutional neural networks for mobile vision applications. Arxiv Preprint Arxiv:1704.04861. https://doi.org/10.48550/arXiv.1704.04861

Hu J, Shen L, and Sun G (2018). Squeeze-and-excitation networks. In the Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, IEEE, Salt Lake City, USA: 7132-7141. https://doi.org/10.1109/CVPR.2018.00745

Huang G, Liu Z, Van Der Maaten L, and Weinberger KQ (2017). Densely connected convolutional networks. In the Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, IEEE, Honolulu, USA: 4700-4708. https://doi.org/10.1109/CVPR.2017.243 **PMCid:PMC5598342**

Li H and Ni J (2024). Proportionate affine projection tanh algorithm and its step-size optimization. Signal Processing, 223: 109553. https://doi.org/10.1016/j.sigpro.2024.109553

Liu B, Wang J, and Li C (2024). Application of PLS–NN model based on mid-infrared spectroscopy in the origin identification of *Cornus officinalis*. RSC Advances, 14(22): 15209-15219. https://doi.org/10.1039/D4RA00953C **PMid:38737973 PMCid:PMC11082643**

Liu L, Wang B, Xu X, and Xu J (2025). A tea classification method based on near infrared spectroscopy (NIRS) and transfer learning. Infrared Physics and Technology, 145: 105713. https://doi.org/10.1016/j.infrared.2025.105713

Lundberg SM and Lee SI (2017). A unified approach to interpreting model predictions. In the 31st Conference on Neural Information Processing Systems, Long Beach, USA: 1-10.

Ma S, He K, and Peng X (2023). Online coal identification based on one-dimensional convolution and its industrial applications. Applied Sciences, 13(17): 9867. https://doi.org/10.3390/app13179867

Martín Andrés A and Álvarez Hernández M (2025). Estimators of various kappa coefficients based on the unbiased estimator of the expected index of agreements. Advances in Data Analysis and Classification, 19(1): 177-207. https://doi.org/10.1007/s11634-024-00581-x

Muhammad W, Bhutto Z, Shah SAR, Shah J, Shaikh MH, Hussain A, Thaheem I, and Ali S (2022). Deep transfer learning CNN based approach for COVID-19 detection. International Journal of Advanced and Applied Sciences, 9(4): 44-52. https://doi.org/10.21833/ijaas.2022.04.006

Neupane D, Kim Y, Seok J, and Hong J (2021). CNN-based fault detection for smart manufacturing. Applied Sciences, 11(24): 11732. https://doi.org/10.3390/app112411732

Ninh DK, Phan KD, Nguyen TTA, Dang MN, Le Thanh N, and Ferrero F (2024). Classification of urea content in fish using absorbance near-infrared spectroscopy and machine learning. Applied Sciences, 14(19): 8586. https://doi.org/10.3390/app14198586

Ong P, Jian J, Li X, Zou C, Yin J, and Ma G (2025). Sugarcane disease recognition through visible and near-infrared spectroscopy using deep learning assisted continuous wavelet transform-based spectrogram. Spectrochimica Acta Part A: Molecular and Biomolecular Spectroscopy, 324: 125001. https://doi.org/10.1016/j.saa.2024.125001 **PMid:39180971**

Promboonruang S and Boonrod T (2023). Deep transfer learning CNN based for classification quality of organic vegetables. International Journal of Advanced and Applied Sciences, 10(12): 203-210. https://doi.org/10.21833/ijaas.2023.12.022

Sen Y, Zhenmin W, Houqing Z, and Wenlong S (2024). Rice variety classification based on optimized near-infrared spectral classification model. Rice Science, 31(1): 6-9. https://doi.org/10.1016/j.rsci.2023.11.003

Shagun S, Bains A, Sridhar K, Dhull SB, Patil S, Gupta VK, Chawla P, and Sharma M (2024). A comprehensive review on impact of post-harvest management and treatment practices on the quality of garlic (*Allium sativum* L) during storage. Scientia Horticulturae, 337: 113586. https://doi.org/10.1016/j.scienta.2024.113586

Shao W, Li Y, Diao S, Jiang J, and Dong R (2017). Rapid classification of Chinese quince (*Chaenomeles speciosa* Nakai) fruit provenance by near-infrared spectroscopy and multivariate calibration. Analytical and Bioanalytical Chemistry, 409: 115-120. https://doi.org/10.1007/s00216-016-9944-7 **PMid:27796451**

Sharifi K, Sheykhi S, and Magami E (2021). First report of garlic leaf blight caused by *Stemphylium vesicarium* in Iran. Journal of Plant Pathology, 103: 1007-1007. https://doi.org/10.1007/s42161-021-00822-4

Shen Y, Yang Z, Khan Z, Liu H, Chen W, and Duan S (2025). Optimization of improved YOLOv8 for precision tomato leaf disease detection in sustainable agriculture. Sensors, 25(5): 1398. https://doi.org/10.3390/s25051398 **PMid:40096213 PMCid:PMC11902794**

Simonyan K and Zisserman A (2014). Very deep convolutional networks for large-scale image recognition. In the 3rd International Conference on Learning Representations, San Diego, USA: 1-14.

Tian F, Wang L, and Xia M (2022). Signals recognition by CNN based on attention mechanism. Electronics, 11(13): 2100. https://doi.org/10.3390/electronics11132100

Tudu CK, Dutta T, Ghorai M et al. (2022). Traditional uses, phytochemistry, pharmacology and toxicology of garlic (*Allium sativum*), a storehouse of diverse phytochemicals: A review of research from the last decade focusing on health and nutritional implications. Frontiers in Nutrition, 9: 949554. https://doi.org/10.3389/fnut.2022.929554 **PMid:36386956 PMCid:PMC9650110**

Vijaykumar KN, Kulkarni S, Hiremath SM et al. (2023). Field screening of garlic genotypes for identification of resistant sources against purple blotch disease. International Journal of Plant and Soil Science, 35(14): 365-370. https://doi.org/10.9734/ijpss/2023/v35i143059

Workneh YY, Legesse NH, Shiferaw HK, and Ashenafi BD (2024). Management of garlic white rot (*Stromatinia cepivora*) with fungicides and host resistance in North Shewa, central highland of Ethiopia. Journal of Plant Pathology, 106(3): 1335-1345. https://doi.org/10.1007/s42161-024-01690-4

Yeboah PN and Baz Musah HB (2022). NLP technique for malware detection using 1D CNN fusion model. Security and Communication Networks, 2022: 2957203. https://doi.org/10.1155/2022/2957203

Zewde T, Fininsa C, Sakhuja PK, and Ahmed S (2007). Association of white rot (*Sclerotium cepivorum*) of garlic with environmental factors and cultural practices in the North Shewa highlands of Ethiopia. Crop Protection, 26(10): 1566-1573. https://doi.org/10.1016/j.cropro.2007.01.007

Zhai C, Wang W, Gao M, Feng X, Zhang S, and Qian C (2024). Rapid classification of rice according to storage duration via near-infrared spectroscopy and machine learning. Talanta Open, 10: 100343. https://doi.org/10.1016/j.talo.2024.100343

Zhang H, Hu L, Liang W, Li Z, Yuan M, Ye Y, Wang Z, Ren Y, and Li X (2024). BCT-OFD: Bridging CNN and transformer via online feature distillation for COVID-19 image recognition. International Journal of Machine Learning and Cybernetics, 15: 2347-2366. https://doi.org/10.1007/s13042-023-02034-x

Zou X, Wang Q, Chen Y et al. (2025). Fusion of convolutional neural network with XGBoost feature extraction for predicting multi-constituents in corn using near infrared spectroscopy. Food Chemistry, 463: 141053. https://doi.org/10.1016/j.foodchem.2024.141053 **PMid:39241414**