

Interpretable deep learning with hyperspectral imaging for Hawthorn cultivar classification and quality prediction

Ruixin Bai^a, Hongpeng Wang^c, Hui Wang^a, Meiqi Luan^a, ZiJian Liu^c, Bin Yang^a, Zihan Zhao^a, Zhilai Zhan^a, Chu Zhang^{b,*}, Jian Yang^{a,*}

^a State Key Laboratory for Quality Ensurance and Sustainable Use of Dao-di Herbs, National Resource Center for Chinese Materia Medica, China Academy of Chinese Medical Sciences, Beijing, 100700, China

^b School of Information Engineering, Huzhou University, Huzhou, 313000, China

^c Zhejiang Provincial Collaborative Innovation Center of Agricultural Biological Resources Biochemical Manufacturing, Zhejiang University of Science and Technology, Hangzhou, 310023, Zhejiang, China

ARTICLE INFO

Keywords:

Hawthorn (*Crataegus pinnatifida*)
Hyperspectral imaging
Deep learning
EfficientNet
Explainable artificial intelligence

ABSTRACT

Hawthorn (*Crataegus pinnatifida*) is a commonly consumed medicinal fruit. This study proposes a rapid and non-destructive technique that integrates hyperspectral imaging (HSI) with interpretable deep learning for the classification of hawthorn cultivars from different regions and the quantitative prediction of key quality indicators, including citric acid, total sugar, and vitamin C content. A total of 1227 samples were collected from 11 categories, representing different cultivars and origins. Model robustness was ensured by acquiring HSI data in three different orientations, with the stalk positioned horizontally, up, and down. Classification results showed an EfficientNet model achieved the highest accuracy (95.92 %) by fusing spectral data from all three orientations. In the regressions, the EfficientNet model outperformed both PLSR and CNN. Among the measured compounds, citric acid and total sugar yielded satisfactory results, with R^2 values of 0.94 and 0.92 and RPD values of 4.08 and 3.55, respectively. Furthermore, combining gradient-weighted class activation mapping (Grad-CAM) and Shapley additive explanations (SHAP) enabled a visual and quantitative interpretation of spectral feature contributions, effectively addressing black-box issues. This is the first study to integrate HSI with interpretable deep learning for simultaneous classification and quality prediction in hawthorn, with improved model performance through multi-orientation spectral fusion.

1. Introduction

Hawthorn (*Crataegus pinnatifida*), a member of the Rosaceae family, produces sweet and sour fruit that can be used as both medicine and food. Hawthorn is rich in organic acids, sugars, vitamin C, flavonoids, and other bioactive compounds, conferring antioxidant, anti-inflammatory, and hypoglycemic properties (Cui et al., 2024; Zhang et al., 2020). These properties have established hawthorn as an important raw material for food, healthcare, and pharmaceutical products (Liu et al., 2024a). China is the world's largest producer of hawthorn, with a cultivation area of ~86,700 hectares and an annual production exceeding 1.5 million tons (Sang et al., 2023). However, the chemical composition of hawthorn varies greatly across cultivars and geographical origins due to genetic and environmental factors (Liu et al., 2023;

Nychas et al., 2017). Key quality indicators, such as the sugar-to-acid ratio and vitamin C content, determine fruit flavor, nutritional value, and processing suitability (Gundogdu et al., 2024). Thus, to protect high-quality local cultivars and prevent economic fraud, it is essential to develop reliable methods for cultivar and origin identification, as well as the quantitative evaluation of quality-related components, including total sugar, organic acids, and vitamin C.

Traditional methods used for identifying hawthorn cultivars and origins, as well as determining chemical components, primarily rely on subjective evaluations or physical and chemical techniques. These include high-performance liquid chromatography (Ornelas-Lim et al., 2021), liquid chromatography-mass spectrometry (Duan et al., 2022; Ye et al., 2021), gas chromatography-mass spectrometry, and DNA barcoding technology (Wu et al., 2021). Although accurate, these

* Corresponding authors at: No.16, Nanxiao street, Dongzhimen, Dongcheng District, Beijing, 100700, China; Chu Zhang: (C. Zhang); Full postal address: No. 759, East 2nd Road, Huzhou, 313000, China.

E-mail addresses: chuzh@zjhu.edu.cn (C. Zhang), yangchem2012@163.com (J. Yang).

approaches require complex sample preparation, rely on expert operation, and are often destructive, time-consuming, and costly, limiting their suitability for the large-scale, real-time monitoring of food quality. In contrast, spectral technologies offer rapid and non-destructive alternatives. Near-infrared (NIR) spectroscopy, combined with chemometrics or traditional machine learning methods, has been applied to several hawthorn studies, including the prediction of chemical composition (Ye et al., 2023), detection of adulteration in hawthorn powder (Sun et al., 2020), and identification of different processed hawthorn products (Gao et al., 2024). However, these methods typically require destructive pretreatment (e.g., drying, grinding) and collect only single-point spectra, thereby restricting efficiency and comprehensive evaluation (Ma et al., 2020).

Hyperspectral imaging (HSI) is an emerging non-destructive evaluation (NDE) technique that provides both rich spectral information and high-accuracy image features from individual samples (Wang et al., 2025). HSI has been increasingly applied for food quality evaluation due to its rapidity, non-destructiveness, and capability of processing multiple samples at once (Das et al., 2025). Hyperspectral data contains significant amounts of information, but with high dimensionality and redundancy. Therefore, conventional hyperspectral analysis typically requires the selection of characteristic wavelengths to reduce dimensionality. Common feature wavelength selection techniques include the successive projections algorithm (SPA), competitive adaptive reweighted sampling (CARS), and principal component analysis (PCA) (Luka et al., 2024). The extracted wavelengths are then used to build classification or regression models based on conventional machine learning algorithms. However, one disadvantage of this strategy is that the selected wavelengths typically vary depending on the dataset, algorithm, and acquisition conditions, resulting in significant inconsistency and model dependency (Zhang et al., 2022a; Zhou et al., 2023).

In contrast, deep learning models, offering strong self-learning and generalization capabilities, are particularly well-suited for processing high-dimensional complex data. These algorithms can learn in layers by combining low-level features to form more abstract higher-level representations. They can also replace manual feature extraction techniques and address the issue of poor model adaptability (Ahmed et al., 2025a). The combination of HSI technology with deep learning methods has been widely reported in both regression and classification tasks involving various fruits (Chen et al., 2025), such as blueberries, citrus, and peanuts (Shanthini et al., 2025; Li et al., 2025; Wang et al., 2024). However, most of these studies have neglected the influence of spectral acquisition direction on model performance. Typically, samples are either placed in a fixed orientation during image acquisition or spectra from multiple views are simply averaged, potentially leading to the loss of direction-specific information. However, integrating spectral data collected from multiple directions can provide more comprehensive sample characteristics. To date, this combination of HSI and deep learning has yet to be applied to the classification of hawthorn origin, cultivar type, or the quantitative evaluation of quality related components.

Although deep learning models provide strong predictive capabilities, they are often trained in an end-to-end manner, where raw spectral-spatial data are directly mapped to outputs without handcrafted feature extraction. However, such networks contain numerous optimized parameters, making them essentially black boxes with limited interpretability, which may hinder trust in their decisions. Explainable AI (XAI) methods, such as Shapley additive explanations (SHAP) and gradient-weighted class activation mapping (Grad-CAM), can mitigate this limitation by identifying important spectral features and clarifying model behavior (Ryo, 2022; Ahmed et al., 2024; Kalopesa et al., 2023). In the process of predicting key chemicals in hawthorn, we hypothesize that XAI may help to clarify the relationship between spectral data and chemical composition, potentially providing a reference for developing more targeted and effective quality control strategies, as well as offering a scientific basis for designing customized spectral instruments.

This study proposes a rapid and non-destructive method that combines HSI in the visible to near-infrared (VNIR) and short-wave infrared (SWIR) ranges with advanced deep learning techniques. The main objectives of this study are to: (1) build a multi-orientation hawthorn HSI dataset simulating real-world collection scenarios; (2) develop classification models to distinguish cultivars and geographical origins using single- and multi-orientation spectra; (3) establish regression models for predicting citric acid, total sugar, and vitamin C; and (4) apply Grad-CAM and SHAP to interpret spectral contributions. This study is the first to use HSI data in the SWIR and VNIR ranges (combined with XAI technology) to classify common hawthorn cultivars from major producing areas in China. It also investigates the quantitative prediction of key quality-related compounds, including citric acid, total sugar, and vitamin C, while fusing spectral data acquired from multiple sample orientations used to enhance model robustness.

2. Materials and methods

2.1. Sample preparation

In this study, a total of 1227 fresh hawthorn samples representing different cultivars were collected from five provinces across China. Cultivar information and corresponding sample sizes are listed as follows. Each cultivar was denoted by a code comprised of two parts: a two-letter abbreviation representing the province of origin (e.g., HB = Hebei, HN = Henan, SD = Shandong, SX = Shanxi, LN = Liaoning), followed by a three-letter abbreviation of the cultivar name derived from its pinyin initials (e.g., DJX = 'Dajinxing,' YRH = 'Yanranghong,' DWL = 'Dawuling'). The number in parentheses indicates the number of samples collected for each cultivar. Specifically, the dataset includes 'Dajinxing' from Qinghe, Hebei (HB_DJX, 116); 'Yanranghong' from Shijiazhuang, Hebei (HB_YRH, 41); 'Tieshanzha' from Xinglong, Hebei (HB_TSZ, 135); 'Dawuling' from Jiyuan, Henan (HN_DWL, 115); 'Jinruyi' from Feixian, Shandong (SD_JRY, 120); 'Yuganhong' from Xintai, Shandong (SD_YGH, 160); 'Dajinxing' from Yuncheng, Shanxi (SX_DJX, 170); 'Dawuling' from Yuncheng, Shanxi (SX_DWL, 50); 'Shuangguoshanzha' from Shanxi (SX_SG, 50); 'Benxi No.4' from Benxi, Liaoning (LN_BX, 110); and 'Liaohong' from Shenyang, Liaoning (LN_LH, 160). All samples were freshly harvested from local orchards to ensure high quality and consistency. Only intact, defect-free fruits were included in the study and each sample was carefully positioned on a black background plate to maintain consistent imaging conditions during HSI data acquisition. Each fruit was then processed to measure citric acid, total sugar, and vitamin C content.

2.2. Data collection using a hyperspectral imaging system

Given the natural shape of hawthorn fruit, HSI data was acquired from three orientations to ensure comprehensive spectral coverage and reduce potential spectral variations caused by different orientations and uneven surface reflectance. As shown in Fig. 1A, position 1 involved placing the fruit horizontally with the stalk facing sideways, while the stalk side faced downward in position 2 and upward in position 3. A hyperspectral imaging system (HySpex VNIR-1800 and HySpex SWIR-384, Norsk Elektro Optikk, Oslo, Norway) spanning both the visible to near-infrared (VNIR, 410–990 nm) and short-wave infrared (SWIR, 940–2500 nm) ranges was used for HSI data acquisition. The system consisted of two HSI cameras, two 150 W halogen lamps (H-LAM, Norsk Elektro Optikk), an electronically controlled translation stage, and a computer for control and data storage. Each sample was placed on a black background and scanned under symmetrical illumination provided by two halogen lamps. The acquisition settings were as follows: exposure times of 35 ms (VNIR) and 45 ms (SWIR), a stage speed of 2.5 mm/s, and an object distance of 320 mm. Black-and-white calibration was applied, to correct for image noise and uneven illumination, using the following formula:

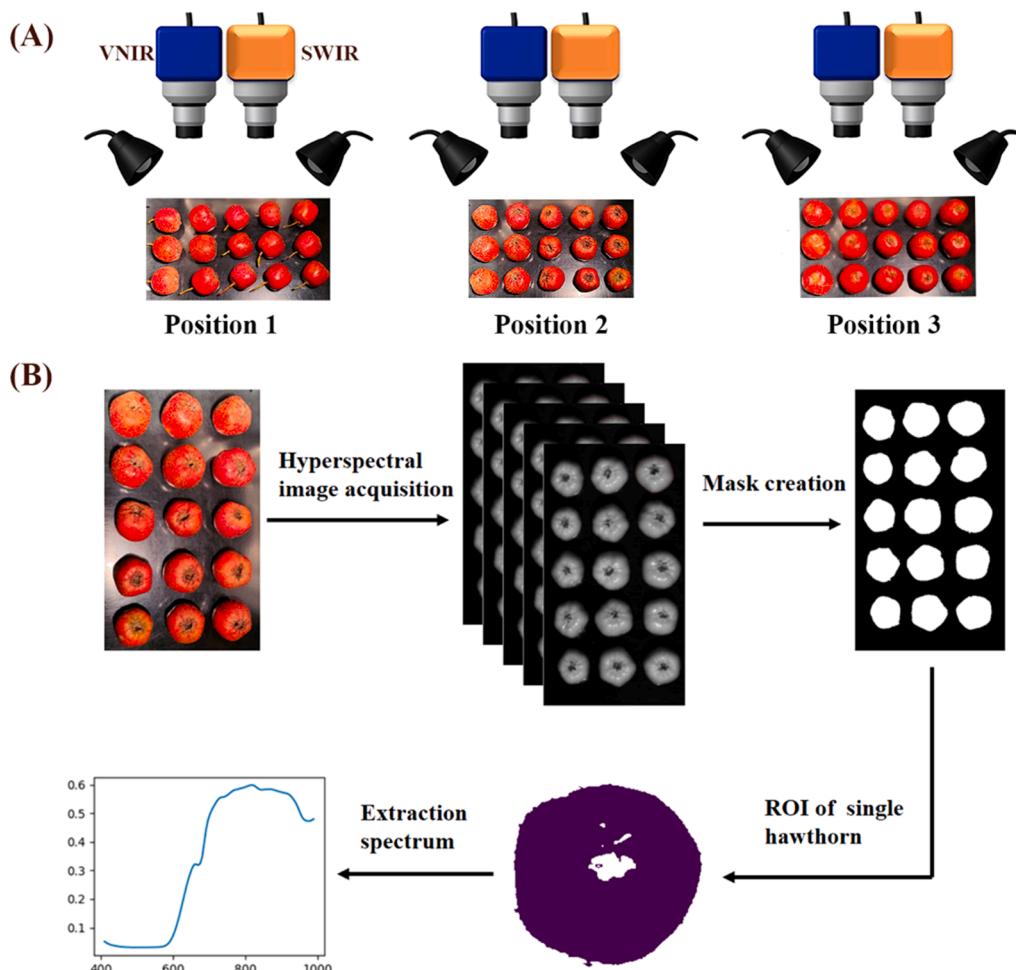


Fig. 1. (A) An illustration of hawthorn sample placement positions during hyperspectral imaging, including three orientations: Position 1 – stalk side facing horizontally; Position 2 – stalk side facing downward; and Position 3 – stalk side facing upward. (B) A schematic diagram of the region of interest (ROI) extraction and spectral data acquisition workflow (using Position 2 as an example).

$$R = \frac{(I_{\text{raw}} - I_{\text{white}})}{(I_{\text{white}} - I_{\text{black}})}$$

where R represents the corrected hyperspectral image, I_{raw} is the original image, I_{white} is the image captured with a white Teflon board, and I_{black} is the image taken with the black lens cover over the camera.

2.3. Region of interest (ROI) extraction

Regions of interest (ROIs) were automatically extracted from hyperspectral images of hawthorn samples using a fixed threshold segmentation method. The original hyperspectral image was first masked to remove background information and isolate the potential sample area, as illustrated in Fig. 1B. Principal component analysis (PCA) was subsequently conducted to reduce data dimensionality and emphasize sample-related spatial features. The first principal component (PC1), which accounted for the majority of spectral variance, was selected for further processing. Specifically, the PC1 image was binarized using an empirically determined threshold (VNIR: -0.75; SWIR: -0.50), wherein pixel values exceeding this threshold were assigned a value of 1 (foreground), and the remainder were set to 0 (background). Morphological operations, including closing and dilation, were then performed to improve segmentation accuracy by eliminating noise and enhancing region continuity. Connected component labeling was also used to assign unique identifiers to each contiguous region, thereby facilitating precise ROI delineation. Finally, the spectral profile of each sample was

obtained by calculating the mean reflectance across all pixels within the corresponding ROI. It should be noted that in position 2 (stalk facing downward), small hollow regions occasionally appeared in the extracted ROI due to the presence of the sepal cavity (Fig S1). These hollows typically accounted for <5 % of the fruit surface and were composed mainly of sepal tissues rather than edible flesh, thus having minimal impact on the representativeness of the extracted spectral information.

2.4. Determination of citric acid, total sugar, and vitamin C content in hawthorn

Chemical composition analysis involved the random selection of 50 samples from each variety, excluding 'Yanranhong' (HB_YRH), for which only 41 samples were available due to limited sample quantity. All selected samples were freeze-dried and ground into a fine powder using a 60-mesh sieve to ensure homogeneity. Total sugar content was determined using a commercial analytical kit (Solarbio, China) and manufacturer instructions. Glucose was used as a calibration standard and the results were expressed as milligrams of glucose per gram weight (mg/g). Citric acid and vitamin C content were measured using ultra-performance liquid chromatography-quadrupole linear ion trap tandem mass spectrometry (UPLC-Q/TRAP-MS/MS). Sample extraction involved 20 mg of freeze-dried hawthorn powder homogenized in 1.5 mL of 80 % methanol (for citric acid determination) or 2 % oxalic acid solution (for vitamin C determination). The mixture was then extracted by ultrasonication for 30 min and centrifugation at 10,000 rpm for 10

min. The resulting supernatant was filtered through a $0.22\text{ }\mu\text{m}$ membrane prior to instrumental analysis. Detection methods for citric acid and vitamin C were conducted using the protocols described by Lu et al. (2017) and Zhou et al. (2019), respectively.

2.5. Deep learning algorithm for hawthorn classification and content prediction

Convolutional neural networks (CNNs) are a type of feed-forward deep learning architecture that perform well in both classification and regression tasks (Saha and Manickavasagan, 2021). CNNs can effectively learn spatial correlation patterns from input data by stacking convolutional layers and nonlinear transformation layers for hierarchical feature extraction, thereby achieving robust analysis of high-dimensional signals (Saha and Manickavasagan, 2021). A typical CNN consists of an input layer, multiple convolutional layers with nonlinear activation functions, pooling layers for dimensionality reduction, and fully connected layers for final prediction (Gao et al., 2018). In this study, model

performance was improved further (while ensuring computational efficiency) by implementing EfficientNet, a CNN class developed using a compound scaling method that can uniformly adjust network depth, width, and resolution. EfficientNet combines a mobile-side inverted bottleneck convolution (MBConv) module with a squeeze-excitation (SE) attention mechanism to adaptively recalibrate channel-level feature responses, significantly improving performance with fewer parameters (Tan and Le, 2019). In this study, CNN and EfficientNet models were developed and optimized with the following primary objectives: (1) classify hawthorn cultivars from different regions and (2) predict the content of key chemical components, including citric acid, total sugar, and vitamin C.

Classification and regression tasks were performed by optimized CNN- and EfficientNet-based models using spectral input data. The CNN-based model performed ~100 training iterations to optimize key hyperparameters, such as the number of convolutional layers, fully connected layer size, learning rate, activation function, and dropout rate. A ReLU activation function was applied after each convolutional

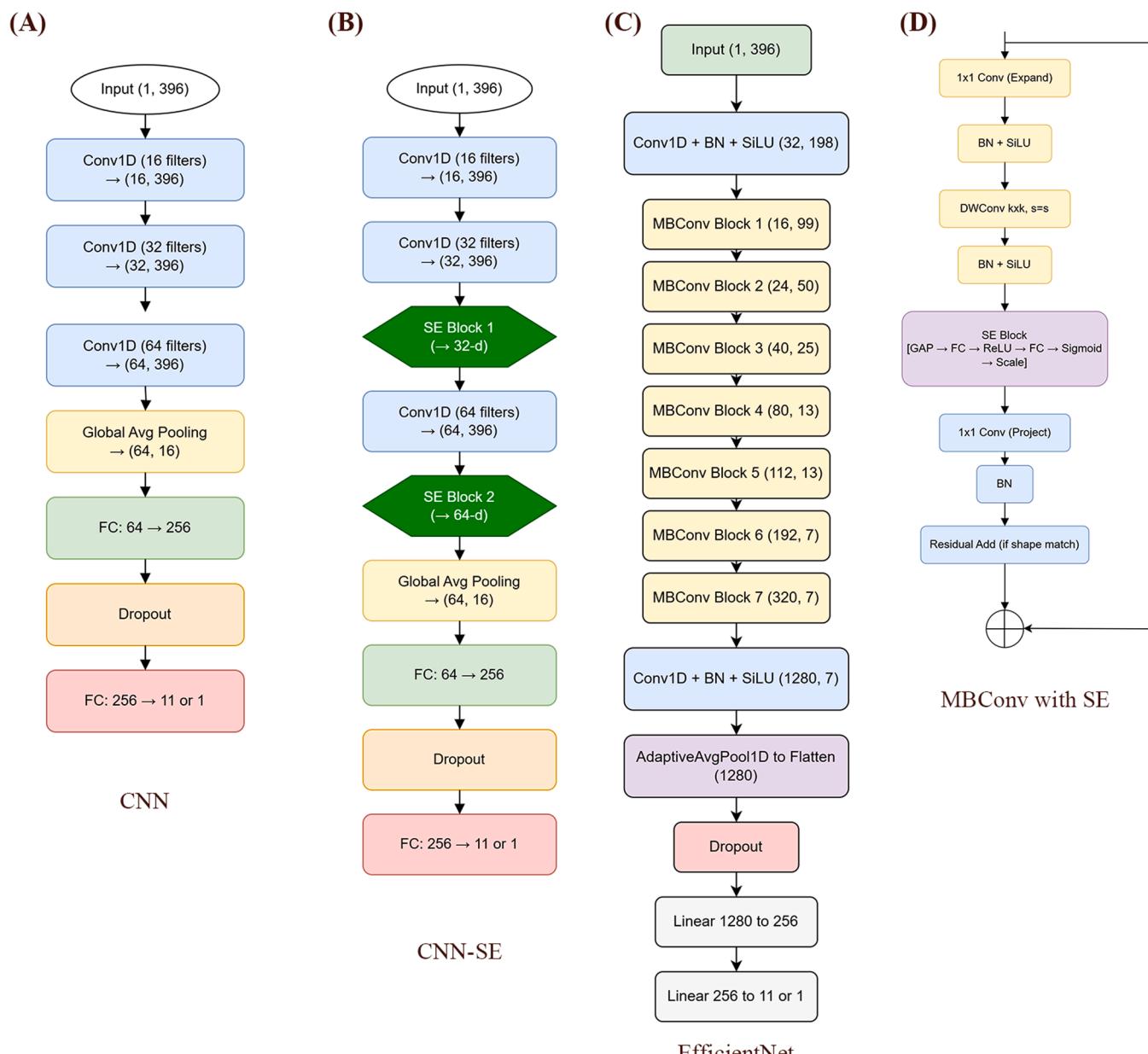


Fig. 2. Deep learning network architectures. (A) The CNN model. (B) The CNN-SE model. (C) The baseline architecture of the EfficientNet model. (D) The internal structure of the MBConv block used in the EfficientNet model.

layer and fully connected layer. A Softmax activation function, optimized using a cross-entropy loss function, was used to generate class probabilities for classification tasks. The output layer, used for regression tasks and optimized by a mean square error (MSE) loss function, consisted of one neuron with a linear activation function. A squeeze and excitation (SE) attention module was also added to the last two convolutional blocks to enhance feature representations. The CNN and CNN-SE network architectures are shown in Fig. 2A-B.

The EfficientNet model used in this study is based on the EfficientNet-B0 architecture, which includes a backbone convolutional layer followed by 7 MBConv modules that employ depthwise separable convolutions to improve parameter efficiency. SE attention modules were also integrated to adaptively recalibrate channel features. A sigmoid linear unit (SiLU) activation function was consistently used throughout the network, due to its smooth and non-monotonic properties, which are highly advantageous for deep neural architectures. Classification tasks involved the EfficientNet model, which employed a global average pooling layer followed by a fully connected layer with a Softmax activation function. This model was then trained using cross-entropy loss to distinguish between 11 hawthorn cultivar categories. In regression tasks, the classification head was replaced by a single linear output node and trained using MSE loss to predict the concentration of citric acid, total sugar, or vitamin C. This model architecture, based on EfficientNet, is shown in Fig. 2C-D. All models were trained using the Adam optimizer with a learning rate of 0.001, to balance convergence speed and training stability. An early stopping strategy based on validation loss was adopted with a patience time of 10 epochs, to alleviate overfitting and enhance generalizability.

2.6. Model interpretability

Model interpretability was improved using two XAI techniques: Grad-CAM and SHAP. These methods provide valuable insights into the decision-making processes of deep learning models by highlighting the most influential features.

2.6.1. Grad-CAM

Grad-CAM is a gradient-based visualization technique that generates class-discriminative localization maps using gradients flowing into selected convolutional layers (van Zyl, Ye, and Naidoo, 2024; Narkhede, 2024). These maps highlight regions in the input that contribute most to model output. In this study, Grad-CAM was adapted for a regression task, enabling the identification of spectral features that are most relevant for predicting continuous chemical composition values, including citric acid, total sugar, and vitamin C concentrations. The stem convolutional layers in the EfficientNet model were selected as target layers for Grad-CAM analysis, to balance interpretability with spatial resolution. This choice ensured the visualization focused on early feature representations that retained rich spectral information.

2.6.2. SHAP

The SHAP method was used to quantitatively explain model predictions, based on initial insights obtained from Grad-CAM. SHAP is derived from cooperative game theory and assigns an additional importance value to each feature, thus providing a method for explaining complex machine learning models (van Zyl, Ye, and Naidoo, 2024; Narkhede, 2024). In this study, SHAP was applied to an EfficientNet-based regression model used to evaluate the contributions of individual spectral bands to predicted concentrations of citric acid, total sugar, and vitamin C. The SHAP method selected KernelExplainer and used a background dataset consisting of 100 randomly selected test samples to approximate local model behavior. SHAP analysis then identified the most influential wavelengths for each target compound. Combining this information with Grad-CAM visualization results facilitated a better understanding of how spectral features drove model predictions, thereby enhancing model interpretability.

2.7. Dataset partitioning

Classification tasks involved a combination of the synthetic minority oversampling technique (SMOTE) and stratified sampling, used to address class imbalance problems and ensure balanced class distributions. The dataset was divided into calibration, validation, and prediction sets using a ratio of 7:2:1. To maintain the integrity of model evaluation, SMOTE was only applied to the calibration set, while the validation and prediction sets remained unchanged. Regression tasks involved randomly dividing the data into calibration, validation, and prediction sets using the same ratio of 7:2:1 without employing any data augmentation techniques.

2.8. Software and model evaluation

HSI data processing and model development were conducted on a Windows 10 platform using Python (3.9) and PyCharm Community. Deep learning models were developed using DL Pytorch. All data analysis steps were performed on a PC equipped with 64 GB RAM, an INTEL i9-10,900 K CPU, and an NVIDIA GEFORCE RTX 4080 GPU (with CUDA version 11.4). An optimal model was produced by using the calibration set for training and employing the validation set to adjust model parameters. Several metrics were used to evaluate model classification performance, including accuracy, recall, F1-score, and area under the receiver operating characteristic (ROC) curve (AUC-ROC). Together, these metrics provided a robust assessment of the model's ability to correctly identify and classify different hawthorn cultivars (You et al., 2024). Their definitions are as follows:

$$\text{Accuracy} = \frac{TP + TN}{TP + TN + FP + FN}$$

$$\text{Recall} = \frac{TP}{TP + FN}$$

$$\text{Precision} = \frac{TP}{TP + FP}$$

$$\text{F1-score} = 2 \cdot \frac{\text{Precision} \cdot \text{Recall}}{\text{Precision} + \text{Recall}}$$

Model performance during regression tasks was evaluated using the determination coefficient (R^2), Pearson correlation coefficient (r), root mean square error (RMSE), mean absolute error (MAE), and performance deviation ratio (RPD) (Ding et al., 2025):

$$R^2 = 1 - \frac{\sum_{i=1}^n (y_i - \hat{y}_i)^2}{\sum_{i=1}^n (y_i - \bar{y})^2}$$

$$r = \frac{\sum_{i=1}^n (y_i - \bar{y})(\hat{y}_i - \bar{\hat{y}})}{\sqrt{\sum_{i=1}^n (y_i - \bar{y})^2 \cdot \sum_{i=1}^n (\hat{y}_i - \bar{\hat{y}})^2}}$$

$$\text{RMSE} = \sqrt{\frac{1}{n} \sum_{i=1}^n (y_i - \hat{y}_i)^2}$$

$$\text{MAE} = \frac{1}{n} \sum_{i=1}^n |y_i - \hat{y}_i|$$

$$\text{RPD} = \frac{\text{SD}}{\text{RMSE}}$$

$$\text{SD} = \sqrt{\frac{1}{n-1} \sum_{i=1}^n (y_i - \bar{y})^2}$$

where y_i and \hat{y}_i represent the reference and predicted values of the i th sample, and \bar{y} and $\bar{\hat{y}}$ represent the mean of all reference values and the

mean of all predicted values, respectively. Specifically, higher R^2 and r values indicate a better model fit, with smaller RMSE and MAE values representing smaller error. The higher the RPD value, the stronger the stability and generalization ability of the model. Generally, $RPD < 1.5$ suggests that the model is unsuitable for prediction; $1.5 \leq RPD < 2.0$ indicates only rough predictive ability, $2.0 \leq RPD < 3.0$ suggests that the model has good prediction performance, and $RPD \geq 3.0$ suggests that the model has excellent prediction performance (Li et al., 2024).

3. Results and discussion

3.1. Spectroscopic analysis

Outlier detection was performed using a statistical method based on the Z-score and a threshold of 3. Two outlier spectra were thus identified and removed from the entire dataset before analysis, one from position 1 and the other from position 3 (see Fig. S2). Fig. 3 shows average reflectance spectra and combined spectra from three independent positions for each hawthorn variety (after the removal of outliers). Although overall spectral trends were consistent across different positions, there were significant differences in reflectance values, which may be due to different chemical compositions or different surface directions. Characteristic spectral peaks were located at ~ 620 , 970 , 1120 , 1310 , 1640 , 1800 , and 2200 nm, with valleys near 670 , 840 , 970 , 1200 , 1480 , and 1930 nm. A reflectance dip near 600 – 620 nm may be attributable to color-related pigments in hawthorn. Notably, the SD_JRY variety, which

is a yellow-fruited hawthorn, exhibited a distinct absorption peak near 600 nm. The valley near 670 nm belongs to the absorption spectra of carotenoids and chlorophylls in plants (Li et al., 2023). A gentle valley near 840 nm may be due to the third overtone of C–H, while the valley near 970 nm may be related to O–H stretching in the second overtone (Zhong et al., 2023). Wavelengths near 1120 , 1200 , and 1310 nm can be assigned to the second overtone of C–H stretching, while wavelengths of 1480 nm can be assigned to the first overtone of O–H stretching (Zhang et al., 2022). In addition, 1640 and 1800 nm were associated with C–H and O–H stretching, respectively (Guo et al., 2023). Water absorption peaks, located at ~ 1930 nm, were a combination of O–H group stretching (Ma et al., 2020). Peaks near 2200 nm were related to a combination band of C–H stretching (Ma et al., 2020). Although the spectral reflectance from various samples was obviously distinct, it was still difficult to directly distinguish hawthorn cultivars based on raw spectral curves alone. As such, advanced data analysis techniques, including CNN and EfficientNet, were needed to better capture differences between cultivars and establish quantitative relationships between spectral features and the content of quality related compounds.

3.2. Determination of citric acid, vitamin C, and total sugar content

The citric acid, vitamin C, and total sugar content for different hawthorn cultivars were analyzed. These results are shown in Fig. 4 and Table S1. A statistical analysis using the Kruskal-Wallis H test revealed significant differences ($p < 0.05$) in citric acid, total sugar, and vitamin C

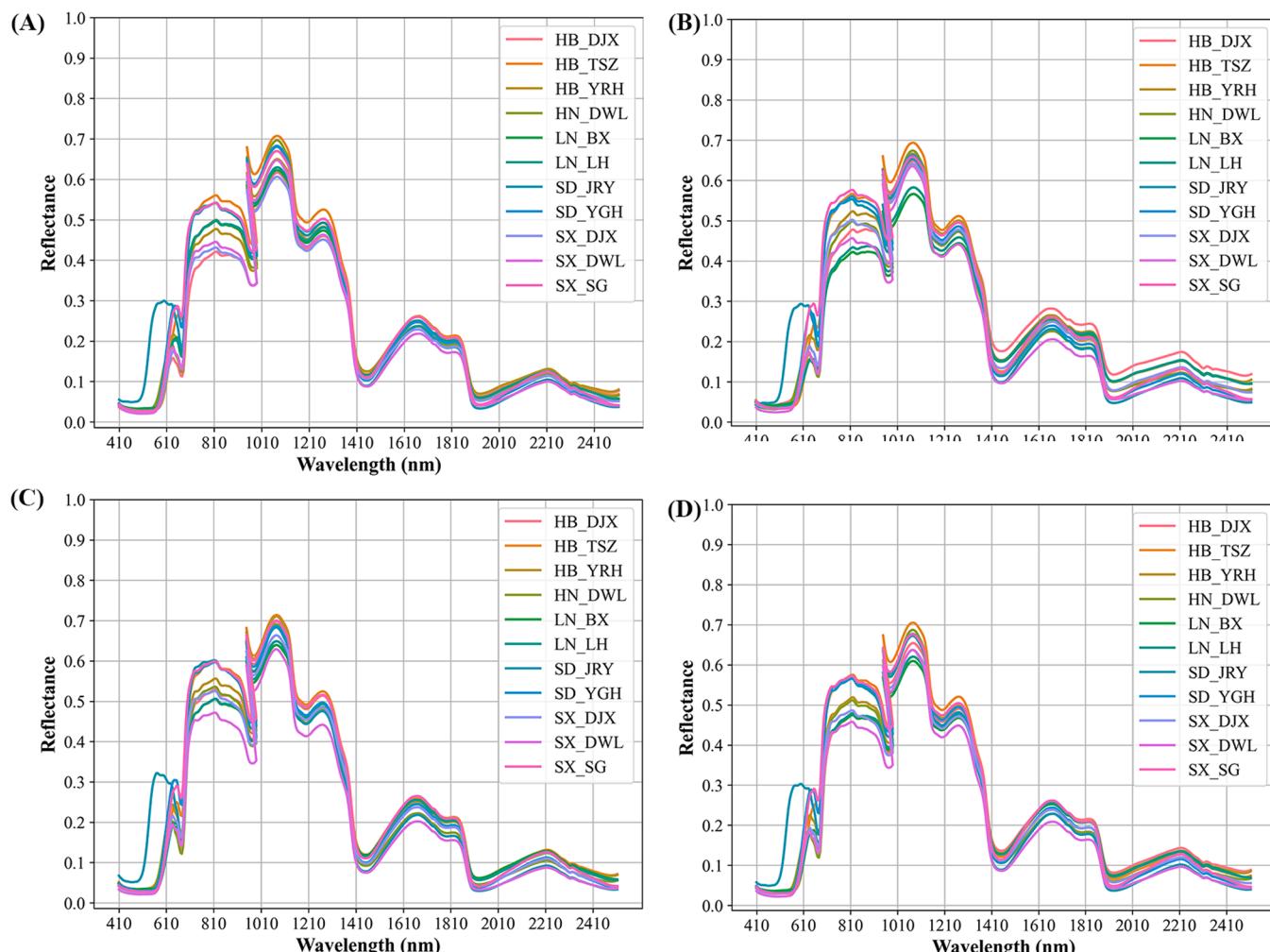


Fig. 3. Average reflectance spectra from 11 hawthorn cultivars at four positions: (A) Position 1; (B) Position 2; (C) Position 3; and (D) combined spectra. Cultivar codes: HB_DJX, HB_YRH, HB_TSZ, HB_DWL, SD_JRY, SD_YGH, SX_DJX, SX_DWL, SX_SG, LN_BX, and LN_LH.

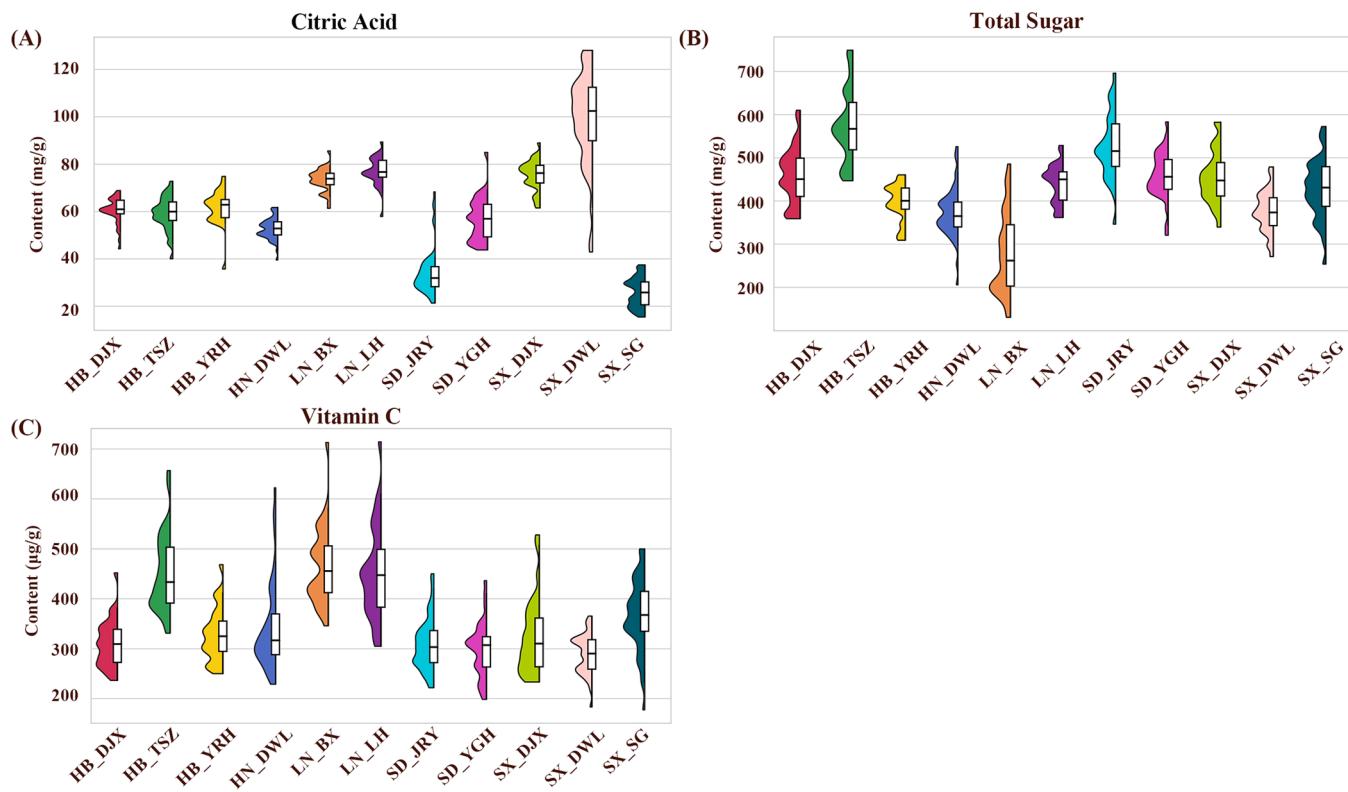


Fig. 4. Combined violin-box plots showing the distribution of (a) citric acid, (b) total sugar, and (c) vitamin C content across different hawthorn cultivars. The violin plots depict the kernel density of the data distribution on one side, while the overlaid box plots display the interquartile range, with the horizontal line indicating the median.

content among different hawthorn cultivars. The SX_DWL cultivar, originating from Shanxi Province, exhibited the highest average citric acid content (96.92 mg/g), indicating a distinctly acidic profile. In contrast, the other two Shanxi cultivars, SX_SG and SD_JRY, showed markedly lower citric acid levels (25.94 mg/g and 34.55 mg/g, respectively). In terms of sweetness, HB_TSZ (from Hebei) and SD_JRY (from Shandong) had the highest total sugar contents, averaging 574.60 mg/g and 514.69 mg/g, respectively, while LN_BX (from Liaoning) exhibited the lowest sugar content (276.79 mg/g). For Vitamin C, the highest levels were observed in HB_TSZ, LN_BX, and LN_LH (also from Liaoning), reaching 453.02 μg/g, 446.15 μg/g, and 435.68 μg/g, respectively, whereas SX_DWL contained the lowest concentration (296.30 μg/g). These differences may be related to genetic background, cultivation practices, or environmental conditions (Gundogdu et al., 2024). Overall, HB_TSZ, SD_JRY, and SX_SG exhibited higher sweetness, making them suitable for fresh consumption. In contrast, SX_DWL had a significantly higher acidity, making it more suitable for processing into sour products such as preserved fruits and jams. These distinct flavor profiles not only provide a reference for consumers when purchasing hawthorn, but also provide a basis for product development tailored to different market needs.

3.3. Classification results for hawthorn cultivars from different origins

The SMOTE algorithm was applied to the calibration set to address the class imbalance problem. In the case of three datasets representing a single orientation, the number of samples in each category increased to 119. In the case of datasets with spectral data fused from three positions, the number of samples in each category increased to 357. The hawthorn dataset was divided using a stratified sampling technique to ensure the proportions of each cultivar in each subset remained consistent. The sample size for each category is listed in Table S2.

3.3.1. Influence of network depth and attention on CNN performance

CNN models with varying depths were evaluated to determine the optimal network configuration, using fused spectral data from three orientations as a representative dataset. Specifically, CNN architectures with three and four convolutional layers were compared. As shown in Tables S3–4, these results indicated that although increasing the number of convolutional layers significantly increased the model complexity, as reflected by the increase in FLOPs and the number of parameters, the prediction accuracy only increased by 0.82 %. This implies that simply deepening the number of network layers did not significantly improve performance, but instead increased the computational burden. The squeeze-excitation (SE) attention mechanism was introduced in the three-layer CNN to further improve performance. This mechanism not only reduced model complexity (FLOPs = 3.36 MMac, parameters = 274.68k), it also improved the test accuracy by 3.95 %, achieving significant improvements in both efficiency and performance. However, when SE was applied to the four-layer CNN, the complexity remained almost unchanged and the accuracy even decreased slightly. Therefore, the CNN models used in subsequent analyses include three convolutional layers.

3.3.2. Impact of fruit orientation and model type on classification performance

This section investigates the impact of fruit orientation and deep learning model architecture on classification performance. As shown in Table 1 and Tables S5–S7, the EfficientNet model consistently outperformed both CNN and CNN-SE models, regardless of hawthorn orientation. In addition, the classification accuracy of spectral data fused from three orientations was significantly higher than that obtained from single orientations, reaching 95.92 %. The CNN-SE model that incorporated data from all three orientations showed several practical advantages, especially in terms of deployment efficiency. Despite its lower complexity, CNN-SE still achieved a classification accuracy of 89.93 %,

Table 1

Comparisons of performance metrics for various models utilizing fused hyperspectral data from all three acquisition positions.

Model	Calibration Set				Validation Set				Prediction Set			
	Recall %	Precision %	F1-score %	Accuracy %	Recall %	Precision %	F1-score %	Accuracy %	Recall %	Precision %	F1-score %	Accuracy %
CNN	88.56	88.59	88.56	88.56	77.58	77.87	77.31	79.29	82.26	79.43	80.48	81.50
CNN-SE	97.58	97.59	97.58	97.58	89.06	90.30	89.39	91.01	89.58	88.43	88.81	89.93
EfficientNet	99.77	99.77	99.77	99.77	93.58	95.74	94.38	95.10	95.42	94.46	94.88	95.92

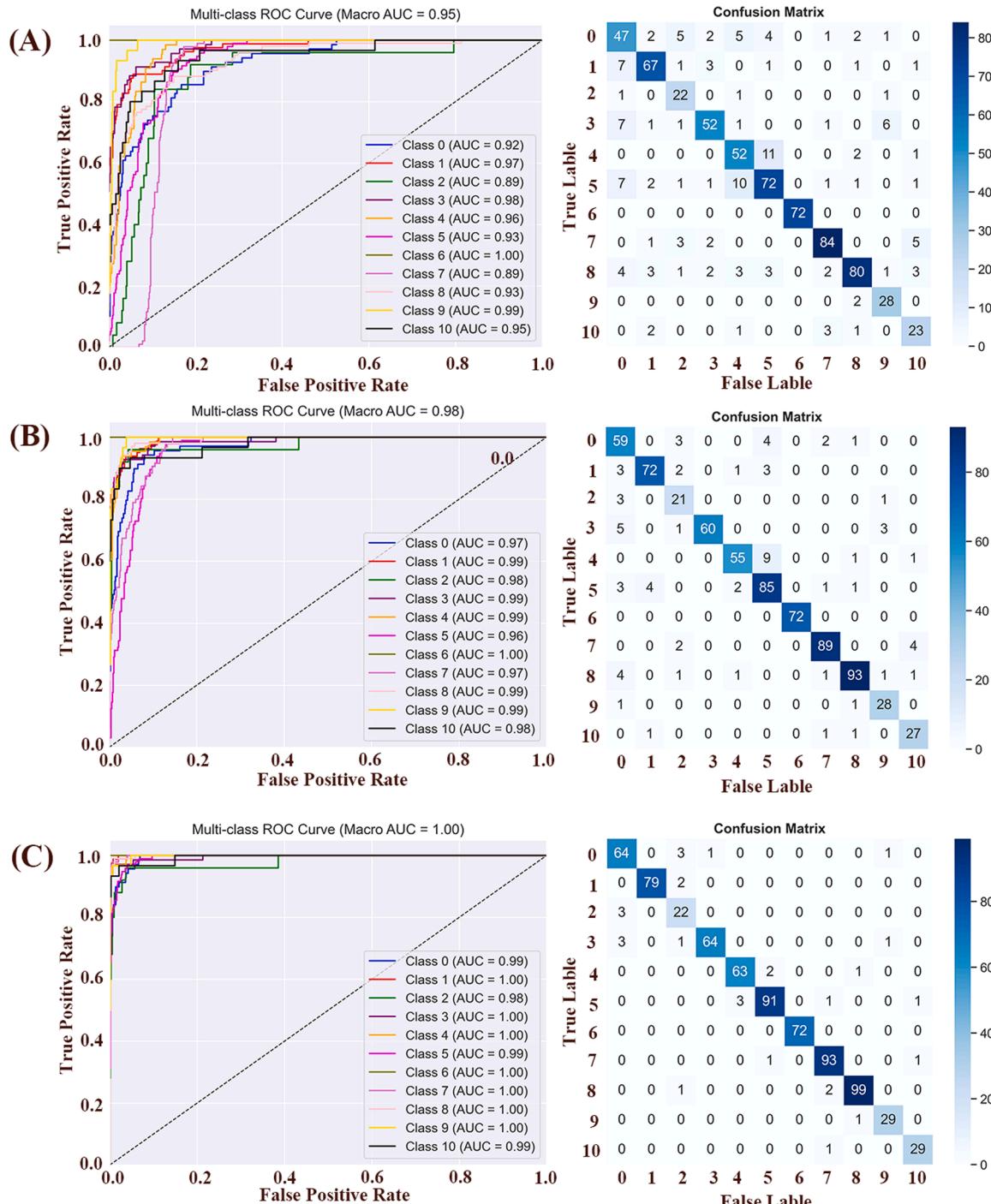


Fig. 5. ROC-AUC curves and confusion matrices for various models using fused hyperspectral spectra from all three acquisition positions. (A) CNN, (B) CNN-SE, (C) EfficientNet. Class labels: 0 = HB_DJX; 1 = HB_TSZ; 2 = HB_YRH; 3 = HN_DWL; 4 = LN_BX; 5 = LN_LH; 6 = SD_JRY; 7 = SD_YGH; 8 = SX_DJX; 9 = SX_DWL; 10 = SX_SG.

demonstrating its potential for practical applications with limited computing resources. For a single orientation, the accuracy of position 1 (stalk facing horizontally) was the lowest among all models, which may be due to limited spectral information available at this angle. In contrast, the spectral data from positions 2 (stalk facing down) and 3 (stalk facing up) were more comprehensive and led to better performance. Specifically, the EfficientNet model achieved an accuracy of 82.80 % at position 1 (Table S5), 93.47 % at position 2 (Table S6), and 93.88 % at position 3 (Table S7). These results further indicate that the small hollow regions occasionally observed in the ROI of position 2 did not compromise classification performance.

Compared to single-direction data, multi-directional fusion improves classification performance for hawthorn cultivars and production areas, highlighting the complementary value of spectral-spatial information. However, in industrial scenarios, sequential acquisition from three directions can increase complexity. In practice, this limitation can be addressed by integrating multiple hyperspectral cameras into a fixed array to simultaneously acquire images from different directions. Furthermore, in future work, we plan to explore advanced modeling strategies to approximate the benefits of multi-directional fusion using only single-view data, such as data-augmented deep learning or uncertainty-guided direction-invariant learning.

3.3.3. ROC curve analysis

The discriminative capabilities of each model were comprehensively evaluated by calculating the area under the ROC curve (AUC) and corresponding confusion matrices. Overall, the EfficientNet models exhibited higher AUC values across all positions. Notably, at position 3, all models achieved AUC scores above 0.96 (see Fig. S5). AUC values further improved when using fused data from all three positions, with all models exceeding 0.98, indicating a significant advantage in overall discriminative capabilities (see Fig. 5). In contrast, the CNN-based models generally exhibited lower AUC values. For example, at position 1, the AUC for class 5 (LN_LH) dropped to 0.85 using the CNN model (see Fig. S3). Similarly, at position 2, the AUC for class 2 (HB_YRH) was only 0.85 for the CNN-SE model, suggesting the model has a limited ability to separate this class from others under certain conditions, such as variations in spectral features or an overlap with other classes (see Fig. S4).

3.3.4. Confusion matrix analysis

Confusion matrices were used to further assess classification performance among hawthorn cultivars. The EfficientNet model maintained consistently high accuracy across 11 classes when using fused spectral data acquired at three positions, achieving a balanced recognition performance (see Fig. 4). At position 3, where the classification accuracy for spectral data from a single position was the highest, classification errors were primarily concentrated in cultivars with fewer samples and highly similar spectral features, resulting in slightly blurred

category boundaries (see Fig. S5). In contrast, at position 1, where the classification accuracy was lower, misclassification occurred for categories 2, 4, 5, and 10, indicating spectral information at this angle may have hindered feature extraction (see Fig. S3). These findings highlight the importance of multi-angle imaging and data fusion for improving model robustness and classification accuracy.

3.4. Regression results for citric acid, vitamin C, and total sugar content

Classification results showed that using fused spectra from three positions provided a clear advantage over single position data. As such, regression analysis was performed using only fused spectra to predict the contents of citric acid, vitamin C, and total sugars.

3.4.1. Regression results

Table 2 summarizes the performance of four regression models (PLSR, CNN, CNN-SE, and EfficientNet), with significant differences observed in the predictive capabilities of each (see Fig. S6–8). For example, the conventional PLSR model exhibited limited performance, with R^2_p values for the three components all below 0.72 and RPD values ranging from 1.31 to 1.82, indicating poor generalizability and limited practical applicability. In contrast, the CNN model and its attention-enhanced variant (CNN-SE) offered improved prediction accuracy for citric acid, with a correlation coefficient (R_p) of 0.88 and an RPD value close to 2, indicating moderate predictive utility. However, in the case of vitamin C and total sugar, CNN and CNN-SE failed to provide significant improvements, with unsatisfactory results for all evaluation metrics, particularly for R^2 values below 0.80 and RPD values of <2.0, indicating limited predictive accuracy in these cases. These results suggest that such models in their current form are not suitable for quantifying these two components.

In contrast, the deep regression model based on EfficientNet showed significant advantages for predicting the content of these three components. In the case of citric acid, the R^2 values from the EfficientNet model applied to the calibration, validation, and prediction sets were 0.93, 0.95, and 0.94, respectively. Corresponding RMSE values were 8.52, 4.83, and 5.34 $\mu\text{g/g}$, respectively, with an RPD of 4.08, indicating the model offers excellent fitting and prediction performance. EfficientNet also produced the best predictions for total sugar, with an R^2_p of 0.92 and an RPD of 3.55. Although the overall performance for all models in vitamin C prediction was relatively low, EfficientNet still achieved acceptable results ($R_p = 0.86$, RPD = 1.93), demonstrating good generalizability. The relatively weak performance for vitamin C may be attributed to its narrow variation range across different collected samples, as well as its high sensitivity to environmental and post-harvest factors (Taghinezhad et al., 2023).

3.4.2. Visualization of prediction results

A scatter plot of the regression fitting curve is provided in Fig. 6,

Table 2

Prediction results for citric acid, total sugar, and Vitamin C content using various regression models.

Parameter	Model	Calibration Set				Validation Set				Prediction Set				
		Rc	RMSEc	R^2_c	MAEc	Rv	RMSEv	R^2_v	MAEv	Rp	RMSEp	R^2_p	MAEp	RPD
Citric Acid	PLSR	0.85	12.45	0.72	8.60	0.79	13.74	0.62	9.23	0.84	11.94	0.70	8.70	1.82
	CNN	0.85	12.71	0.70	8.76	0.81	13.35	0.65	9.04	0.88	10.77	0.76	7.98	2.02
	CNN-SE	0.84	13.24	0.68	9.21	0.81	13.36	0.65	9.45	0.88	11.10	0.74	8.36	1.96
	EfficientNet	0.93	8.52	0.93	3.97	0.98	4.83	0.95	3.10	0.97	5.34	0.94	3.43	4.08
Total Sugar	PLSR	0.80	55.84	0.64	44.24	0.71	67.54	0.49	52.64	0.77	51.62	0.59	40.20	1.56
	CNN	0.65	72.14	0.41	56.50	0.65	73.41	0.40	56.91	0.74	58.25	0.48	45.59	1.39
	CNN-SE	0.67	70.36	0.43	53.88	0.68	70.28	0.45	55.47	0.73	55.29	0.53	42.72	1.46
	EfficientNet	0.95	29.61	0.90	17.94	0.96	26.61	0.92	16.47	0.96	22.76	0.92	14.50	3.55
Vitamin C	PLSR	0.75	58.95	0.56	45.85	0.70	60.92	0.48	48.63	0.66	68.93	0.42	51.00	1.31
	CNN	0.64	68.78	0.40	53.70	0.69	62.04	0.46	49.82	0.59	73.85	0.33	57.48	1.23
	CNN-SE	0.69	65.92	0.45	52.44	0.71	61.17	0.47	48.68	0.63	71.23	0.38	56.23	1.27
	EfficientNet	0.90	40.50	0.79	25.36	0.90	37.23	0.80	23.12	0.86	47.00	0.73	25.76	1.93

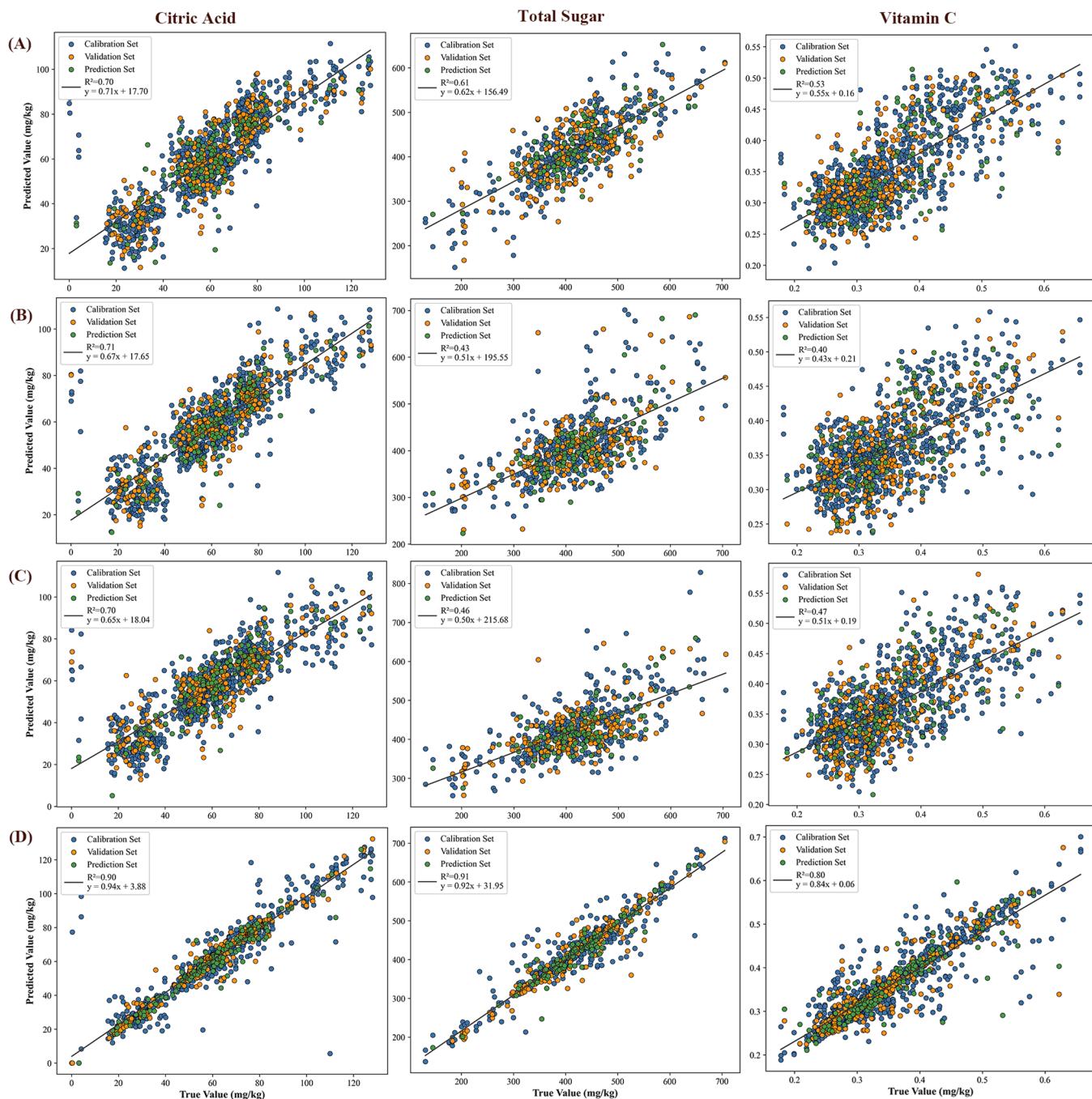


Fig. 6. A comparison of predicted versus actual values using different models for citric acid, total sugar, and vitamin C predictions. Results are shown for (A) PLSR, (B) CNN, (C) CNN-SE, and (D) EfficientNet.

demonstrating the correlation between the predicted and actual values and more intuitively demonstrating the performance. It is evident the EfficientNet regression fitting results for the three target components exhibit a good linear trend. Most of the data points for citric acid and total sugar are densely distributed near the 45° diagonal, indicating the predicted value is highly consistent with the actual value and there are no systematic deviations in the model. In contrast, the traditional CNN and PLSR models showed a certain degree of discreteness. The resulting distribution is relatively scattered, particularly in the high concentration areas where errors are large. Although the point distributions for vitamin C predictions are slightly scattered, EfficientNet still maintained high correlation, indicating the model offers strong robustness for regression tasks with large differences in component content and high

noise levels.

The present study demonstrated that multi-orientation fusion improves classification accuracy and quantitative prediction performance, confirming the complementary value of spectral-spatial information from different views. Nevertheless, sequential acquisition from three orientations may introduce additional complexity in industrial scenarios. In practice, this limitation can be addressed by integrating multiple hyperspectral cameras in a fixed array to capture different orientations simultaneously; an approach that has already been applied in fruit sorting and grading lines. Moreover, advanced modeling strategies, such as orientation-invariant deep learning with data augmentation or uncertainty-guided acquisition, may approximate the benefits of multi-orientation fusion using only single-view data. These findings,

therefore, not only validate the robustness of multi-orientation HSI analysis, but they also provide a theoretical foundation for practical solutions that balance accuracy and efficiency in real-world applications.

3.5. Model interpretability analysis

Since the EfficientNet model showed excellent performance for chemical composition prediction and regression tasks, the Grad-CAM and SHAP methods were employed to study the corresponding spectral response mechanisms and enhance the interpretability of hawthorn sample predictions. As shown in Fig. 7, Grad-CAM highlighted spectral regions to which the model paid more attention during feature extraction, while SHAP values revealed the relative contributions of each wavelength to model output.

3.5.1. Model interpretability analysis for citric acid

In the case of citric acid, Grad-CAM visualization showed the model response was enhanced in the spectral regions of 550–750 nm, 1350–1450 nm, 1850–1950 nm, and 2100–2300 nm, indicating these regions were given priority during feature extraction. An analysis of SHAP values further revealed the most influential wavelengths included 681 nm, 654 nm, 1625 nm, 703 nm, 1385 nm, 1712 nm, 2093 nm, and 2480 nm. Citric acid contains multiple carboxyl groups (-COOH), which exhibit overtones and combination band absorption in the near-infrared (NIR) region, particularly at 1350–1450 nm (O-H bending and second overtone), 1700–1750 nm and 1850–1950 nm (C = O stretching and COOH combination), and 2100–2300 nm (C-O and O-H combination bands) (Kalaimani et al., 2019; Tsioptsias et al., 2024). These absorption features agreed well with the significant spectral regions, indicating the model successfully captured features associated with the carboxylic acid groups. In addition, bands in the visible range (e.g., 681 nm) may reflect weak responses associated with changes in color or reflectance, caused by changes in the state of citric acid.

3.5.2. Model interpretability analysis for total sugars

During the prediction of total sugars, Grad-CAM analysis showed the model paid special attention to spectral regions near 550–700 nm, 1130 nm, 1380–1420 nm, and 1850–1920 nm. SHAP values also highlighted the importance of bands such as 649 nm, 670 nm, 1472 nm, 1734 nm, 1308 nm, 2208 nm, 1701 nm, and 1499 nm. Sugars in hawthorn are primarily composed of glucose, fructose, and sucrose, which are rich in hydroxyl (-OH) functional groups. In addition, their first and second harmonic absorptions are mainly located in the 1130–1250 nm and 1350–1500 nm regions, respectively. The C-H stretching vibration combination bands typically appear between 1700 and 1800 nm and 2100–2300 nm (Bouley, Schulte, and Pierson, 2015). These chemically meaningful bands are well matched to emphasized regions, indicating the model can effectively identify spectral patterns associated with hydroxyl and hydrocarbon groups.

3.5.3. Model interpretability analysis for vitamin C

In the case of vitamin C, the Grad-CAM method exhibited strong responses in the 600–700 nm, 1360–1450 nm, and 1850–1900 nm regions. These responses correspond to key wavelengths identified by SHAP, such as 1505 nm, 1194 nm, 976 nm, 1946 nm, and 960 nm. The molecular structure of vitamin C contains multiple phenolic and alcoholic hydroxyl groups and conjugated C = C double bonds. These features produce distinct absorption bands at 1350–1450 nm (second O-H overtone), 1850–1950 nm (C = C and O-H combination), and 1200 nm (first O-H overtone) (Weyer and Lo, 2006). These characteristic bands closely match regions of interest in the model. In addition, vitamin C may also affect the color and structure of tissues, indirectly affecting the reflectance of visible light in the 600–700 nm range, due to its strong reducing properties.

4. Conclusion

This study proposed a non-destructive, interpretable, deep learning

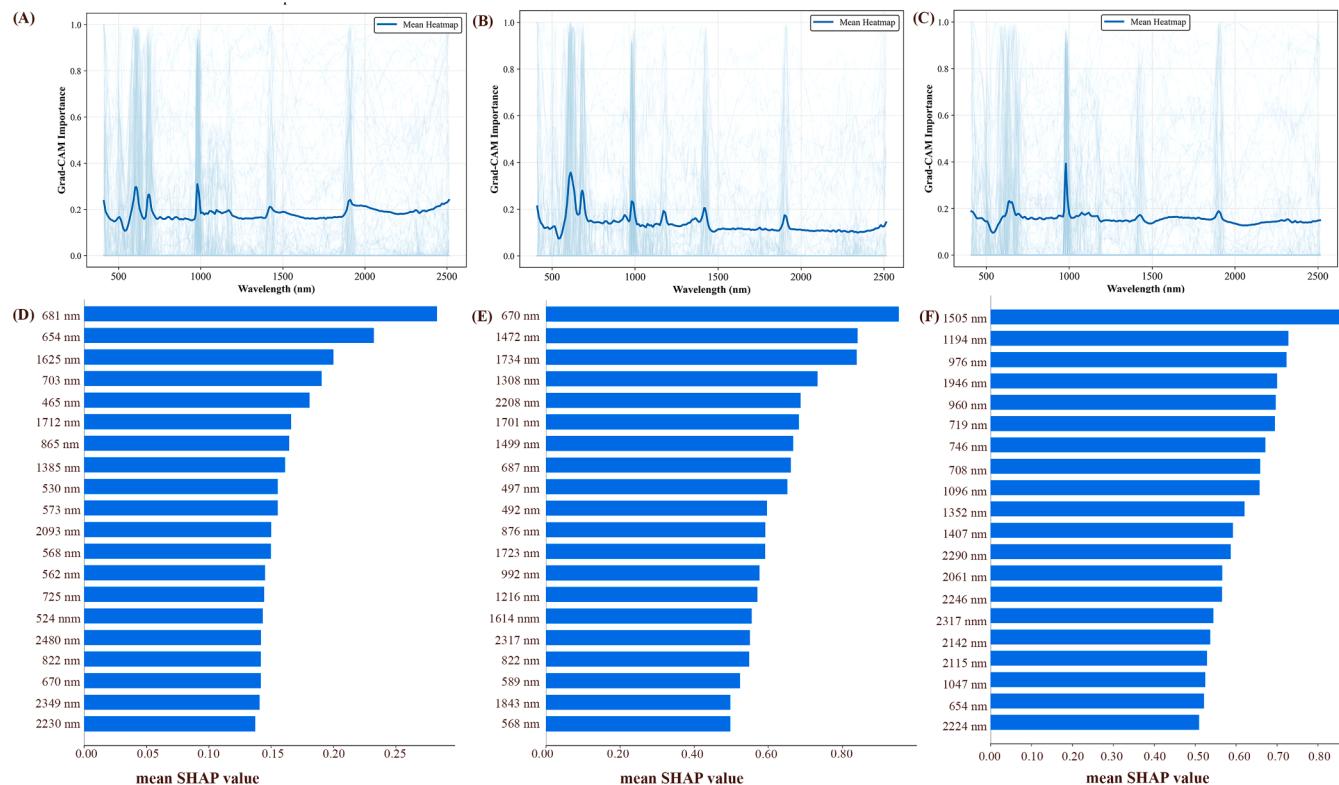


Fig. 7. Grad-CAM and SHAP-based spectral importance plots for the prediction of citric acid (A, D), total sugars (B, E), and vitamin C (C, F) content using the EfficientNet model.

framework based on HSI for the classification of hawthorn (*Crataegus pinnatifida*) cultivars from different origins, as well as the quantitative prediction of key quality-related components (i.e., citric acid, total sugars, and vitamin C). A total of 1227 samples representing 11 hawthorn cultivar categories from different producing areas were collected. Classification and regression models were then developed based on CNN and EfficientNet architectures. This study is the first to integrate XAI with HSI technology across both VNIR and SWIR bands, combining hawthorn spectral data from three different orientations to achieve variety classification and the quantitative prediction of key components. Experimental results showed that fusing spectral data acquired from three different positions significantly improved classification performance, with the EfficientNet model achieving the highest accuracy (95.92%). In scenarios requiring lightweight deployment, the CNN-SE model served as a practical alternative, reaching an accuracy of 89.93% with only 0.27 million parameters. In regression tasks, the EfficientNet model consistently outperformed CNN and traditional PLSR methods, achieving R^2 values of 0.97, 0.96, and 0.86 for citric acid, total sugar, and vitamin C, respectively. In addition, the integration of Grad-CAM and SHAP highlighted spectral regions emphasized by the model and quantified the contribution of each wavelength to individual predictions, offering valuable insights into chemically significant regions, thereby enhancing model transparency and credibility. Overall, the proposed HSI method, combined with deep learning, shows great potential for fast, accurate, and interpretable hawthorn quality assessment. Future research will focus on expanding the dataset to encompass a wider range of cultivars, varying maturity stages, and more diverse environmental conditions, thereby further enhancing the robustness and generalizability of the proposed framework.

Data availability

The processed spectral dataset (extracted from ROIs) and source code used in this study are openly available in Zenodo at <https://doi.org/10.5281/zenodo.17060526>.

Ethical statement - studies in humans and animals

This study did not involve any experiments on human participants or animals.

CRediT authorship contribution statement

Ruibin Bai: Writing – review & editing, Writing – original draft, Software, Conceptualization. **Hongpeng Wang:** Visualization, Software, Methodology. **Hui Wang:** Visualization, Conceptualization. **Meiqi Luan:** Methodology, Formal analysis. **ZiJian Liu:** Methodology. **Bin Yang:** Investigation, Funding acquisition. **Zihan Zhao:** Methodology. **Zhilai Zhan:** Resources, Funding acquisition. **Chu Zhang:** Writing – review & editing, Software. **Jian Yang:** Writing – review & editing, Project administration, Methodology, Funding acquisition, Conceptualization.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgements

This work was supported by the National Key R&D Program of China (2024YFC3506800), the Shandong Provincial Key Research and Development Program (Major Technological Innovation Project) (2021CXGC010508), the Scientific and Technological Innovation Project of China Academy of Chinese Medical Sciences (CI2023E002,

CI2021B009), the National Administration of Traditional Chinese Medicine High-level Key Discipline Construction Project of Traditional Chinese Medicine (ZYYZDXK-2023244), China Agricultural Research System of MOF and MARA (CARS-21), and the Fundamental Research Funds for the Central Public Welfare Research Institutes (ZZ16-YQ-040, ZZXT2023012). We thank LetPub (www letpub.com.cn) for linguistic assistance and pre-submission expert review.

Supplementary materials

Supplementary material associated with this article can be found, in the online version, at [doi:10.1016/j.fufo.2025.100761](https://doi.org/10.1016/j.fufo.2025.100761).

References

- Ahmed, T., Wijewardane, N.K., Lu, Y., Jones, D.S., Kudanov, M., Williams, C., Villordon, A., Kamruzzaman, M., 2024. Advancing sweetpotato quality assessment with hyperspectral imaging and explainable artificial intelligence. *Comput. Electron. Agric.* 108, 108855. <https://doi.org/10.1016/j.compag.2024.108855>.
- Ahmed, M.T., Villordon, A., Kamruzzaman, M., 2025. Hyperspectral imaging and explainable deep learning for non-destructive quality prediction of sweetpotato. *Postharvest Biol. Technol.* 222, 113379. <https://doi.org/10.1016/j.postharvbio.2024.113379>.
- Bouley, L., Schulte, H., Pierson, J.-F., 2015. Near-infrared optical spectroscopy of sugar-based mixtures: a snapshot to identify issues of influence. In: Proceedings of OCM 2015 –2nd International Conference on Optical Characterization of Materials. Karlsruhe, Germany. <https://doi.org/10.58895/ksp/1000044906-3>.
- Cui, M., Cheng, L., Zhou, Z., Zhang, Y., Wang, X., Liu, J., 2024. Traditional uses, phytochemistry, pharmacology, and safety concerns of hawthorn (*Crataegus genus*): a comprehensive review. *J. Ethnopharmacol.* 319, 117229. <https://doi.org/10.1016/j.jep.2023.117229>.
- Chen, Y., Guo, Z., Barbin, D.F., Dai, Z., Watson, N., Povey, M., Zou, X., 2025. Hyperspectral imaging and deep learning for quality and safety inspection of fruits and vegetables: a review. *J. Agric. Food Chem.* 73 (17), 10019–10035. <https://doi.org/10.1021/acs.jafc.4c11492>.
- Das, M., Yeo, W.S., Saptoro, A., 2025. A review of machine learning in hyperspectral imaging for food safety. *Vib. Spectrosc.* 139, 103828. <https://doi.org/10.1016/j.vibspec.2025.103828>.
- Ding, F., García-Martín, J.F., Zhang, L., Xu, Z., Lv, D., Chen, X., Tu, K., Lan, W., Pan, L., 2025. Prediction of quality traits in packaged mango by NIR spectroscopy. *Food Res. Int.* 205, 115963. <https://doi.org/10.1016/j.foodres.2025.115963>.
- Duan, L., Xiong, H., Du, Y., Zhang, Y., Zhou, X., 2022. High-throughput LC-MS method for the rapid characterisation and comparative analysis of multiple ingredients of four hawthorn leaf extracts. *Phytochem. Anal.* 33, 635–643. <https://doi.org/10.1002/pea.3116>.
- Gao, L., Zhong, L., Huang, R., Yue, J., Li, L., Nie, L., Wu, A., Huang, S., Yang, C., Cao, G., Meng, Z., 2024. Identification and determination of different processed products and their extracts of *Crataegi Fructus* by infrared spectroscopy combined with two-dimensional correlation analysis. *Spectrochim. Acta A Mol. Biomol. Spectrosc.*, 123922 <https://doi.org/10.1016/j.saa.2024.123922>.
- Gao, Q., Lim, S., Jia, X., 2018. Hyperspectral image classification using convolutional neural networks and multiple feature learning. *Remote Sens.* 10, 299. <https://doi.org/10.3390/rs10020299>.
- Guo, Z., Zhang, J., Dong, H.W., Liu, R., Zhao, M., 2023. Spatio-temporal distribution patterns and quantitative detection of aflatoxin B1 and total aflatoxin in peanut kernels explored by short-wave infrared hyperspectral imaging. *Food Chem.* 424, 136441. <https://doi.org/10.1016/j.foodchem.2023.136441>.
- Gundogdu, M., Ozrenk, K., Ercisli, S., Hegedus, A., Turan, M., 2024. Organic acids, sugars, vitamin C content and some pomological characteristics of eleven hawthorn species (*Crataegus spp.*) from Turkey. *Biol. Res.* 47, 21. <https://doi.org/10.1186/0717-6287-47-21>.
- Kalaimani, N., Ramya, K., Vinitha, G., Praveen, C., 2019. Structural, spectral, thermal and nonlinear optical analysis of anhydrous citric acid crystal. *Optik* 192, 162960. <https://doi.org/10.1016/j.jleo.2019.162960>.
- Kalopesa, E., Karyotis, K., Tziolas, N., Tsakiridis, N., Samarinis, N., Zalidis, G., 2023. Estimation of sugar content in wine grapes via *in situ* VNIR–SWIR point spectroscopy using explainable artificial intelligence techniques. *Sensors* 23 (3), 1065. <https://doi.org/10.3390/s23031065>.
- Li, C., Jin, C., Zhai, Y., Pu, Y., Qi, H., Zhang, C., 2025. Simultaneous detection of citrus internal quality attributes using near-infrared spectroscopy and hyperspectral imaging with multi-task deep learning and instrumental transfer learning. *Food Chem.* 481, 143996. <https://doi.org/10.1016/j.foodchem.2025.143996>.
- Li, P., Tang, S., Chen, S., Zhang, Y., Wang, H., 2023. Hyperspectral imaging combined with convolutional neural network for accurately detecting adulteration in *Atlantic salmon*. *Food Control*. 147, 109573. <https://doi.org/10.1016/j.foodcont.2022.109573>.
- Li, D., Park, B., Kang, R., Chen, Q., Ouyang, Q., 2024. Quantitative prediction and visualization of matcha color physicochemical indicators using hyperspectral microscope imaging technology. *Food Control* 163, 110531. <https://doi.org/10.1016/j.foodcont.2024.110531>.
- Liu, J., Meng, C., Zhang, K., Wang, Y., Li, F., 2023. Origin traceability of *Yimucao* (Chinese motherwort) in China using stable isotopes and extracts assisted by

- machine learning techniques. *J. Food Compos. Anal.* 126, 105900. <https://doi.org/10.1016/j.jfca.2023.105900>.
- Liu, Z., Gu, J., Zhou, C., Wang, Y., Yang, J., Huang, J., Wang, H., Bai, R., 2024. Identification of geographical origin for hawthorn based on hyperspectral imaging technology. *Sci. Technol. Food Ind.* 45, 282–291. <https://doi.org/10.13386/j.issn1002-0306.2023090074>.
- Lu, C.G., Yang, J., Kang, C.Z., Wang, X.Y., Li, Y., 2017. Determination of nine organic acids in *dendrobium officinale* by UPLC-MS/MS. *J. Chin. Med. Mater.* 40, 1360–1363. <https://doi.org/10.13863/j.issn1001-4454.2017.06.027>.
- Luka, B.S., Yunusa, B.M., Vihikwagh, Q.M., Kuhwa, K.F., Oluwasegun, T.H., Ogalagu, R., Yuguda, T.K., Adnouini, M., 2024. Hyperspectral imaging systems for rapid assessment of moisture and chromaticity of foods undergoing drying: principles, applications, challenges, and future trends. *Comput. Electron. Agric.* 224, 109101. <https://doi.org/10.1016/j.compag.2024.109101>.
- Ma, T., Tsuchikawa, S., Inagaki, T., 2020. Rapid and non-destructive seed viability prediction using near-infrared hyperspectral imaging coupled with a deep learning approach. *Comput. Electron. Agric.* 177, 105712. <https://doi.org/10.1016/j.compag.2020.105712>.
- Narkhede, J., 2024. Comparative evaluation of post-hoc explainability methods in AI: LIME, SHAP, and Grad-CAM. In: Proceedings of the 4th International Conference on Sustainable Expert Systems (ICSES). IEEE, pp. 826–830. <https://doi.org/10.1109/ICSES63445.2024.10762963>.
- Nychas, G.-J., Yin, W., Zhang, C., He, J., Li, X., 2017. Application of near-infrared hyperspectral imaging to discriminate different geographical origins of Chinese wolfberries. *PLoS ONE* 12, e0180534. <https://doi.org/10.1371/journal.pone.0180534>.
- Ornelas-Lim, C., Luna-Vázquez, F.J., Rojas-Molina, A., López-Hernández, J.F., Hernández-Brenes, C., 2021. Development of a quantified herbal extract of hawthorn (*Crataegus mexicana*) leaves with vasodilator effect. *Saudi Pharm. J.* 29, 1338–1345. <https://doi.org/10.1016/j.jps.2021.10.002>.
- Ryo, M., 2022. Explainable artificial intelligence and interpretable machine learning for agricultural data analysis. *Artif. Intell. Agric.* 6, 257–265. <https://doi.org/10.1016/j.aiia.2022.05.001>.
- Saha, D., Manickavasagan, A., 2021. Machine learning techniques for analysis of hyperspectral images to determine quality of food products: a review. *Curr. Res. Food Sci.* 4, 28–44. <https://doi.org/10.1016/j.crfc.2021.01.002>.
- Sang, Y., Hao, Z., Dai, S., Liu, Y., 2023. Optimization of aqueous two-phase extraction of polysaccharides from hawthorn by response surface methodology coupled with genetic algorithm and its antioxidant activity. *Starch/Stärke* 75, 2300094. <https://doi.org/10.1002/star.202300094>.
- Shanthini, K.S., George, S.N., Chandran, O.V.A., Jinumol, K.M., Keerthana, P., Francis, J., George, S., 2025. NorBlueNet: hyperspectral imaging-based hybrid CNN-transformer model for non-destructive SSC analysis in Norwegian wild blueberries. *Comput. Electron. Agric.* 235, 110340. <https://doi.org/10.1016/j.compag.2025.110340>.
- Sun, X., Li, H., Yi, Y., Hua, H., Guan, Y., Chen, C., 2020. Rapid detection and quantification of adulteration in Chinese hawthorn fruits powder by near-infrared spectroscopy combined with chemometrics. *Spectrochim. Acta A Mol. Biomol. Spectrosc.* 250, 119346. <https://doi.org/10.1016/j.saa.2020.119346>.
- Taghinezhad, E., et al., 2023. Predicting quality properties of pears during storage using vis/NIR spectroscopy: impact of storage duration and temperature. *Agriculture* 13 (10), 1913. <https://doi.org/10.3390/agriculture13101913>.
- Tan, M., Le, Q.V., 2019. EfficientNet: rethinking model scaling for convolutional neural networks. In: Daumé III, H., Singh, A. (Eds.), Proceedings of the 36th International Conference on Machine Learning. PMLR, pp. 6105–6114. In: <https://proceedings.mlr.press/v97/tan19a.html>.
- Tsioptsias, C., Panagiotou, A., Mitlianga, P., 2024. Thermal behavior and infrared absorbance bands of citric acid. *Appl. Sci.* 14, 8406. <https://doi.org/10.3390/app14188406>.
- van Zyl, C., Ye, X., Naidoo, R., 2024. Harnessing explainable artificial intelligence for feature selection in time series energy forecasting: a comparative analysis of Grad-CAM and SHAP. *Appl. Energy* 353, 122079. <https://doi.org/10.1016/j.apenergy.2023.122079>.
- Wang, X., Feng, Y., Wang, Y., Li, J., Zhao, H., 2025. Enhancing optical non-destructive methods for food quality and safety assessments with machine learning techniques: a survey. *J. Agric. Food Res.* 19, 101734. <https://doi.org/10.1016/j.jafr.2025.101734>.
- Wang, Z., Yu, Y., Liu, J.Q., Li, X., Ma, H., 2024. Peanut origin traceability: a hybrid neural network combining an electronic nose system and a hyperspectral system. *Food Chem.* 447, 138915. <https://doi.org/10.1016/j.foodchem.2024.138915>.
- Weyer, L.G., Lo, S.-C., 2006. Spectra-structure correlations in the near-infrared. Sample Characterization and Spectral Data Processing. John Wiley & Sons, Ltd., pp. 4102–4109. <https://doi.org/10.1002/0470027320.s4102>.
- Wu, L., Cui, Y., Wang, Q., Zhang, L., Li, X., 2021. Identification and phylogenetic analysis of five *Crataegus* species (Rosaceae) based on complete chloroplast genomes. *Planta* 254, 14. <https://doi.org/10.1007/s00425-021-03667-4>.
- Ye, L.H., Yang, J., Liu, F.M., Chen, J.X., Wang, H., 2021. Ultrahigh-performance liquid chromatography coupled with ion mobility quadrupole time-of-flight mass spectrometry for separation and identification of hawthorn fruits by multivariate analysis. *Microchem. J.* 171, 106832. <https://doi.org/10.1016/j.microc.2021.106832>.
- Ye, T., Zheng, Y., Guan, Y., Sun, Y., Chen, C., 2023. Rapid determination of chemical components and antioxidant activity of the fruit of *Crataegus pinnatifida* Bunge by NIRS and chemometrics. *Spectrochim. Acta A Mol. Biomol. Spectrosc.* 289, 122215. <https://doi.org/10.1016/j.saa.2023.122215>.
- You, S., Li, Y., Song, J., Yu, X., Tu, K., Lan, W., Pan, L., 2024. Evaluating the microstructure and physicochemical properties of 'Korla' fragrant pear disease caused by *Alternaria alternata*: vis-NIR hyperspectral microscope imaging coupled with convolutional neural network. *Postharvest Biol. Technol.* 212, 112913. <https://doi.org/10.1016/j.postharvbio.2024.112913>.
- Zhang, L., Zhang, Q., Wu, J., Liu, Y., Yu, L., Chen, Y., 2022a. Moisture detection of single corn seed based on hyperspectral imaging and deep learning. *Infrared Phys. Technol.* 125, 104137. <https://doi.org/10.1016/j.infrared.2022.104137>.
- Zhang, F., Sun, W., Kong, J., Wang, Y., Li, S., 2022. Detection of moisture content in salted sea cucumbers by hyperspectral and low field nuclear magnetic resonance based on deep learning network framework. *Food Res. Int.* 156, 111174. <https://doi.org/10.1016/j.foodres.2022.111174>.
- Zhang, S.-Y., Sun, X.-L., Yang, X.-L., Li, J.-Y., Chen, Y.-F., Zhao, Q., 2020. Botany, traditional uses, phytochemistry and pharmacological activity of *Crataegus pinnatifida* (Chinese hawthorn): a review. *J. Pharm. Pharmacol.* 74, 1507–1545. <https://doi.org/10.1093/jpp/rga050>.
- Zhong, Q., Zhang, H., Tang, S., Chen, Y., Liu, M., 2023. Feasibility study of combining hyperspectral imaging with deep learning for chestnut-quality detection. *Foods* 12, 2089. <https://doi.org/10.3390/foods12102089>.
- Zhou, H., Yan, L., Hong, Z., 2023. Quantitative aflatoxin B1 detection and mining key wavelengths based on deep learning and hyperspectral imaging in subpixel level. *Comput. Electron. Agric.* 206, 107561. <https://doi.org/10.1016/j.compag.2023.107561>.
- Zhou, L., Guo, L., Gao, F., Zhang, H., Li, X., 2019. Determination of 6 water-soluble vitamins in *Polygoni Multiflori Radix* from different habitants by UPLC-MS/MS. *World Chin. Med.* 14, 1673–1680. <https://doi.org/10.3969/j.issn.1673-7202.2019.11.008>.