# Synergizing Wood Science and Interpretable Artificial Intelligence: Detection and Classification of Wood Species Through Hyperspectral Imaging

**Yicong Qi [1], Yin Zhang [1], Shuqi Tang [2,\*] and Zhen Zeng [1,3]**

[1] College of Architecture, Southwest Minzu University, Chengdu 610041, China; cdzhangyin@163.com (Y.Z.)

[2] College of Engineering, South China Agricultural University, Guangzhou 510642, China

[3] Sichuan Forestry Survey, Design and Research Institute Company Limited, Chengdu 620363, China

\* Correspondence: sqtang@stu.scau.edu.cn

**Abstract:** With the increasing demand for wood in the wood market and the frequent trade of high-value wood, the accurate identification of wood varieties has become essential. This study employs two hyperspectral imaging systems—visible and near-infrared spectroscopy (VNIR) and short-wave infrared spectroscopy (SWIR)—in combination with a deep learning model to propose a method for wood species identification. Spectral data from wood samples were obtained through hyperspectral imaging technology, and classification was performed using a combination of convolutional neural networks (CNNs) and Transformer models. Multiple spectral preprocessing and feature extraction techniques were applied to enhance data quality and model performance. The experimental results show that the full-band modeling is significantly better than the feature-band modeling in terms of classification accuracy and robustness. Among them, the classification accuracy of SWIR reaches 100%, the number of model parameters is 1,286,228, the total size of the model is 4.93 MB, and the Floating Point Operations (FLOPs) is 1.29 M. Additionally, the Shapley Additive Explanation (SHAP) technique was utilized for model interpretability, revealing key spectral bands and feature regions that the model emphasizes during classification. Compared with other models, CNN-Transformer is more effective in capturing the key features. This method provides an efficient and reliable tool for the wood industry, particularly in wood processing and trade, offering broad application potential and significant economic benefits.

**Keywords:** wood; hyperspectral imaging; deep learning; explainable AI; visualization

## 1. Introduction

Wood, as an important natural resource, is widely used in various industries, including construction, furniture, and paper manufacturing [1]. In China's long-standing architectural history, wood has also been extensively used in the construction and restoration of ancient buildings, becoming an essential component of cultural heritage preservation. Golden Phoebe, Hemlock, Cypress, and Camphor Pine are the most common representative woods used in ancient architecture and high-end furniture. These woods are not only favored for their unique grain and high market value, but also widely used for their excellent mechanical and durability properties [2]. For example, Golden Phoebe was used for palace beams and high-end furniture, Hemlock for structural framing and bridges,

Cypress for outdoor architecture and landscaping, and Camphor Pine for homes and light structures. These woods are significant in the modern marketplace and in the preservation of cultural heritage. With the growing market demand and the increase in illegal behavior, the phenomenon of timber counterfeiting, wrong use, and improper substitution is becoming more and more serious. This not only interferes with the material's traceability for ancient building repairs, but also may lead to structural safety hazards, quality problems, and even loss of cultural values [3]. Meanwhile, wood in ancient buildings undergoes significant changes in its properties due to long-term weathering, further increasing the complexity of restoration. Given these challenges, accurate identification of wood species has become critical [4]. However, traditional visual and physical inspection methods are difficult to meet modern market requirements as they are too dependent on experience [5]. In order to solve these problems, the development of efficient and accurate wood identification technology has become an urgent issue for the industry [6].

Wood species identification methods based on spectral technology are non-destructive, efficient, and accurate, making them increasingly important in the modern wood market [7]. Spectral technology captures subtle differences in the chemical composition and physical structure of various wood types, offering a more precise and objective identification method compared to traditional techniques. Commonly used spectroscopic techniques include near-infrared spectroscopy (NIR), mid-infrared spectroscopy (MIR), and Raman spectroscopy, all of which analyze and classify samples based on their chemical fingerprint and molecular properties [8–10]. However, most existing research focuses on NIR technology. For instance, some researchers have used NIR for the classification and evaluation of waste wood, density testing, etc., [7,11,12]. Despite the success of NIR technology in wood analysis, its limited band range and resolution make it difficult to differentiate between wood species with similar physical properties, and measurements are made only for specific monitoring points. In contrast, hyperspectral imaging (HSI) technology collects continuous spectral data over a broader range and provides higher resolution within a narrow wavelength range. HSI generates a three-dimensional data cube containing both spatial and spectral information, which can simultaneously present the physical structure and chemical composition of wood [13,14]. This capability gives HSI an advantage over NIR in analyzing wood's internal composition and structure. HSI is widely used in agriculture, food science, and large-scale remote sensing missions due to its ability to capture both spatial and spectral information. For example, Feng et al. used a self-designed convolutional neural network (CNN) as the base network of a deep migration learning method combined with hyperspectral data for rice disease detection [15]. Other studies use HSI data to analyze food quality, such as detecting adulteration in salmon and detecting antibiotic content in mutton [13,16]. In the field of remote sensing, Yadav et al. proposed a novel pixel deterministic active learning (PCAL) method based on extended differential pattern (EDP) for land cover classification, which significantly improves the classification accuracy and Kappa coefficient through distributed intensity filtering (DIF) and histogram equalization (HE) preprocessing [17]. These studies have demonstrated the potential of hyperspectral imaging techniques for extracting high-dimensional spectral features and spatial information in different fields. However, most of the current studies have focused on specific spectral features, and existing methods have limitations in dealing with objects with complex textures or small spectral differences. In contrast, wood science poses higher demands for feature extraction and classification due to the complex structure and diverse composition of its species. This makes it challenging for existing methods to identify wood species, and there is still relatively little research addressing these unique issues, and the potential of hyperspectral imaging in this area has yet to be fully exploited.

Deep learning, particularly CNN, automatically extracts key features of spectral data through convolution operations. It is good at processing spectral data and has strong generalization ability [18]. However, CNN faces limitations in capturing global features, especially when dealing with complex spatial and spectral information. To address this issue, the Transformer architecture has emerged as a promising approach for wood species recognition. Originally designed for natural language processing, the Transformer's self-attention mechanism has also proven highly effective in image analysis [19]. Combining CNN with Transformer leverages the local feature extraction capability of CNN and the global feature modeling strength of Transformer, which can improve the accuracy and efficiency of recognition. Zheng et al. proposed a real-time dual-branch multi-scale wood recognition network. The network is based on extracting local and global information from images using CNN and transformer branches. The local and global information is dynamically extracted in an input-dependent manner and then interacted through a token mixing module, achieving good results [20]. However, compared to hyperspectral data, RGB images lack detailed spectral information, which is crucial for capturing the subtle chemical and physical differences between wood species. Zhao et al. proposed a cable insulation defect classification method that combines ultrasonic cable insulation defect detection with a CNN-Transformer. The method adopts a CNN-before-Transformer architecture, using CNN to extract local features and then modeling global features through Transformer. Although this method has achieved certain results in detecting and classifying cable defects, the initial extraction of local features may ignore the global dependencies in high-dimensional data, resulting in limited global modeling effects and failure to fully capture cross-dimensional feature interactions [21]. At the same time, interpretability has become a crucial focus in machine learning research [22]. With the rapid development of artificial intelligence, model transparency and interpretability are increasingly important. In this context, Shapley Additive Explanations (SHAP), a model-agnostic game theory-based method, has become a key tool in explainable AI. SHAP generates feature contributions for each instance, facilitating the analysis of how different input features affect recognition results, particularly in hyperspectral data processing, thereby improving model transparency and credibility [23]. However, since SHAP is a relatively recent introduction, its application in spectral data analysis for interpreting predictive models is less common.

To achieve progress, this study aims to (1) apply hyperspectral imaging technology in combination with traditional machine learning algorithms and CNN-Transformer models to identify wood species by collecting spectral data from samples; (2) compare full-band and feature-band modeling to verify the advantages of full-band modeling in classification performance, demonstrating that retaining all spectral band information improves accuracy and robustness; and (3) use SHAP technology for visualization analysis, revealing key spectral bands and feature areas that the model focuses on during classification, thus enhancing interpretability and providing a basis for further optimization.

## 2. Materials and Methods

### 2.1. Sample Preparation

This study focuses on four types of wood: Golden Phoebe, Hemlock, Cypress, and Camphor Pine wood. For each type, 400 wood blocks were prepared with dimensions of 4 cm × 4 cm × 3 cm, resulting in a total of 1600 experimental samples. The wood samples were sourced from different regions in China: Golden Phoebe from Ya'an City, Sichuan Province; Hemlock from Shunchang County, Fujian Province; Cypress from Shifang County, Sichuan Province; and Camphor Pine from the Greater Khingan Range in Heilongjiang Province. The wood types are illustrated in Figure 1. To mitigate the effects of saw marks and aging on the wood surfaces, all blocks were sanded using 100-grit

sandpaper (particle size 150 μm). The samples were then stored in a controlled environment at a temperature of (20 ± 2) °C and a relative humidity of (60 ± 2) % to ensure consistency in experimental conditions.



**Figure 1.** Sample species, Golden Phoebe, Hemlock, Cypress, and Camphor Pine in that order.

*2.2. Hyperspectral Imaging System Spectra Acquisition*

2.2.1. Hyperspectral Imaging System

Two hyperspectral imaging systems were utilized: one for visible and near-infrared hyperspectral imaging (VNIR-HSI, 400–1000 nm) and another for shortwave infrared hyperspectral imaging (SWIR-HSI, 900–1700 nm). As depicted in Figure 2, both systems operate from a shared workstation. The setup includes a dark chamber, two hyperspectral cameras (Specim FX10 and Specim FX17, manufactured by Spectral Imaging Ltd. in Oulu, Finland), a set of 280W halogen lamps (DECOSTAR 51S, produced by Osram Corp. in Munich, Germany), a motorized stage (HXY-OFX01, supplied by Hongxing Yang Technology Co., Ltd., Wuhan, China), and the Lumo-Scanner image acquisition software. The parameter settings for the two imaging systems are shown in Table 1.



**Figure 2.** Diagram of VNIR and SWIR Hyperspectral Imaging Systems (VNIR: 400–1000 nm, SWIR: 900–1700 nm).

**Table 1.** Parameter Settings for VNIR-HSI and SWIR-HSI Systems.

| Parameter | VNIR-HSI System | SWIR-HSI System |
| --- | --- | --- |
| Movement Speed | 9.8 mm/s | 7.5 mm/s |
| Spectrometer Exposure Time | 2 ms | 3.2 ms |
| Distance Between Lens and Sample | 32 cm | 32 cm |
| Spectral Range | 400–1000 nm | 900–1700 nm |
| Average Spectral Interval | 2.68 nm | 1.67nm |

2.2.2. Spectra Extraction and Data Split

Following the acquisition of hyperspectral images, the spectral data were extracted from the regions of interest (ROIs) within each wood sample. To eliminate variations in illumination and sensor noise in the images, the raw spectral data were first subjected to

black-and-white reference correction ($I_B$, $I_W$). The correction process was conducted according to Equation (1):

$$R_c = \frac{R_{raw} - I_B}{I_W - I_B} \tag{1}$$

where $R_c$ is the corrected reflectance image, and $R_{raw}$ is the raw reflectance image.

After calibration, the corrected spectral data were extracted from the ROI of each sample. The dataset was then divided using the Sample set Partitioning based on joint X-Y distances (SPXY) algorithm. Specifically, the 1600 samples were split into a training set and a test set at a 7:3 ratio, resulting in 1120 samples for training and 480 samples for testing. The SPXY algorithm ensures that both sets are well-balanced in terms of sample distribution and response variables by calculating joint distances in both the feature and response spaces, thereby enhancing the model's performance [24].

### 2.3. Spectral Preprocessing

During hyperspectral image acquisition, factors such as illumination conditions, instrument performance, and sample surface inhomogeneity often cause baseline drift and nonlinear scattering effects, which in turn lead to degradation of model performance [25,26]. In order to improve the data quality and enhance the model performance, a variety of preprocessing methods were applied to the raw spectra in this study, including Savitzky–Golay Smoothing (SG), Normalization (NOR), Baseline Correction (BL), Standard Normal Variate (SNV), and Multiplicative Scatter Correction (MSC). SG is used to smooth the spectral signal and reduce the interference of random noise; NOR standardizes the spectral data to a uniform range to eliminate the deviation caused by intensity differences between samples; BL corrects the baseline drift of the spectral signal to ensure the consistency of the spectral curve; SNV enhances the recognition of spectral features by correcting the scattering effect; MSC is used to adjust the spectral intensity and reduce the deviation caused by the unevenness of the sample surface. These methods work together to effectively improve the quality and reliability of the data and provide better input data for the classification model.

### 2.4. Feature Wavelength Selection

Feature selection in spectral data analysis is an effective strategy for mitigating the challenges of dimensionality reduction and addressing multicollinearity in spectral models [27]. This study employs three feature selection methods: Competitive Adaptive Reweighted Sampling (CARS), Successive Projections Algorithm (SPA), and Recursive Feature Elimination (RFE), to evaluate which features most effectively improve the performance of classification models, considering the possibility that feature extraction may not always yield optimal features.

CARS is a feature selection method that combines Monte Carlo sampling with the regression coefficients of a PLS (Partial Least Squares) model, emulating the principle of 'survival of the fittest' from Darwin's theory [28]. The method operates as follows: (1) Monte Carlo sampling is used to randomly select a subset of wavelengths from the full spectrum to construct the PLSR model; (2) an exponential decay function is applied to reduce the variable space, followed by adaptive reweighted sampling to retain variables with larger absolute regression coefficients as a new subset; and (3) in multiple iterations, the wavelength subset with the minimum root mean square error of cross-validation (RMSECV) is chosen as the optimal set.

SPA is a forward wavelength selection algorithm that begins with a single wavelength and incrementally adds a new variable in each iteration until reaching the desired number of selected variables, denoted as N. The primary objective of SPA is to select

wavelengths with the least redundancy in spectral information, thereby addressing the problem of collinearity [29].

RFE is an iterative feature selection method. The core idea is to iteratively train the model and gradually remove features that contribute the least, ultimately selecting the features that most enhance model performance [30]. The process is as follows: (1) Train a base model—in this study, a decision tree—and calculate the importance coefficient w for each feature. (2) Remove the features with the lowest importance coefficients. (3) Repeat these steps until the remaining features reach a predetermined number. (4) Through continuous and recursive elimination, RFE retains the features that most significantly improve the model's effectiveness.

*2.5. Modeling Algorithm and Model Evaluation*

2.5.1. Modeling Algorithm

Partial Least Squares Discriminant Analysis (PLS-DA) is a supervised dimensionality reduction and classification method that has been widely used in the field of spectral classification in recent years. It is based on the principle of separating variable data and categorical information into two datasets by maximizing the variance between sample categories while minimizing the variance within categories [31]. While dimensionality reduction, PLS-DA retains important features that can best distinguish between categories, and therefore can effectively highlight inter-group differences, allowing samples from different categories to be clearly separated.

Extreme Gradient Boosting (XGBoost) is a highly efficient tree-boosting algorithm that delivers fast and accurate solutions to a wide range of data analysis problems. Built on the gradient boosting framework, XGBoost iteratively constructs decision trees to minimize the loss function by focusing on the gradient of the loss rather than just residuals [32]. Additionally, XGBoost incorporates both L1 and L2 regularization in its objective function, which helps mitigate the risk of overfitting. In classification tasks, each tree contributes a real-valued score to the final prediction, and the combined scores determine the final class label.

1D-CNN is particularly used in spectral data analysis. 1D-CNN, a variant of CNN, is specifically designed to process sequential data, such as time series, signals, and spectral data. At its core lies a one-dimensional convolutional layer, which extracts features by sliding a convolutional kernel over a single dimension of the data. In addition, 1D-CNN includes a pooling layer, a fully connected layer and a normalization layer, with the pooling layer helping to reduce feature dimensionality, reduce computation, and provide translation invariance. However, although CNN performs well in extracting local features, it may have limitations in dealing with long-range dependencies. For spectral data, interrelationships and long-distance dependencies between multiple wavelengths are important, which requires a method that can better capture these global features [33,34].

The self-attention mechanism of the Transformer encoder was introduced to establish direct connections between different wavelengths, enabling a more comprehensive understanding of spectral features. The self-attention mechanism is formulated as follows:

$$Attention(Q, K, V) = softmax\left(\frac{QK^T}{\sqrt{d_k}}\right)V \tag{2}$$

In spectral data classification, the query matrix $Q$, the key matrix $K$, and the value matrix $V$, represent feature representations at different wavelengths. This mechanism calculates the similarity between each wavelength $Q$ and all other wavelengths $K$, converting it into weights using the *softmax* function, *softmax* is used in this mechanism to normalize the similarity between wavelengths into a probability distribution, thereby generating weights that highlight the characteristics of high-correlation wavelengths and

weaken the influence of low-correlation wavelengths, ensuring that the model can effectively capture key spectral information to improve classification performance. These weights are then applied to the value matrix $V$ generating a weighted feature representation. In this study, the combination of Transformer encoder and 1D-CNN enables the model to capture both global dependencies and local features in the spectral data. The Transformer encoder first processes the spectral data and extracts the global features, followed by 1D-CNN's local refinement of these features, which further improves the accuracy and robustness of the classification. The specific model structure is shown in Figure 3.
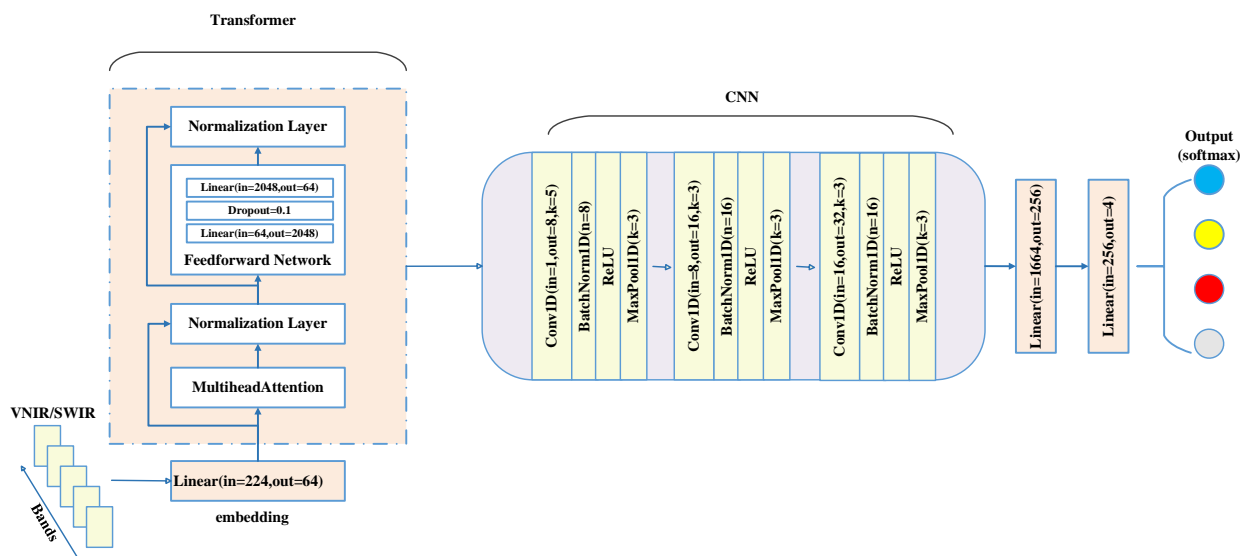


**Figure 3.** CNN-Transformer structure.

### 2.5.2. Model Evaluation

To evaluate the reliability and stability of the model, we calculated the accuracy, precision, recall, and F1-Score. Accuracy, as defined in Equation (3), represents the proportion of correctly identified samples to the total number of samples. Precision, defined in Equation (4), measures the model's ability to correctly predict positive samples. Recall, given by Equation (5), reflects the model's capability to correctly identify positive instances among all actual positive cases. The F1-Score, described by Equation (6), is the harmonic mean of precision and recall, providing a comprehensive metric that balances these two aspects.

$$Accuracy = \frac{TP + TN}{TP + FN + FP + TN} \tag{3}$$

$$Precision = \frac{TP}{TP + FP} \tag{4}$$

$$Recall = \frac{TP}{TP + FN} \tag{5}$$

$$F1 - Score = \frac{2 * Precision \times Recall}{Precision + Recall} \tag{6}$$

In these equations, True Positive (TP) refers to the number of positive samples correctly identified as positive. False Positive (FP) denotes the number of negative samples incorrectly classified as positive, and False Negative (FN) indicates the number of positive samples mistakenly identified as negative. In addition, we also use parameters, model

size, and Floating Point Operations (FLOPs) to evaluate model performance: parameters refers to the total number of trainable parameters of the model, reflecting the complexity of the model; model size refers to the storage space occupied by the model, measuring its deployment efficiency; FLOPs is used to evaluate the computational complexity of the model, representing the amount of computation required for each inference. These indicators comprehensively reflect the efficiency and usability of the model.

### 2.6. Model Interpretation

In interpretable AI, the Shapley Additive Explanation (SHAP) value is widely recognized as a key tool for measuring the importance of features [35]. The SHAP value is based on the theoretical framework of Cooperative Game Theory, and provides a way of quantifying the contribution of each feature to the model's predictions. By analyzing the impact of features in different combinations, the SHAP value can accurately calculate the independent contribution of each feature to the prediction results, while taking into account the interactions between features. Using SHAP values, the importance of input features can be ranked and a feature importance graph can be generated. In such charts, features with higher SHAP values are assigned higher rankings, indicating that these features have a greater influence on model predictions; whereas features with lower SHAP values reflect their lesser influence. Equation (7) quantifies the contribution of each input feature to the results of the data instance.

$$S_{SHAP_x}(v) = \sum_{V:v \in V} \left(|V| \times \begin{matrix} v \\ |V| \end{matrix}\right)^{-1} \left(\widehat{x_V} - \hat{x}_{V \setminus v}\right) \tag{7}$$

Where $v$ is a feature, $x$ is a data instance, $S_{SHAP_x}(v)$ is the computed SHAP value of $v$ for $x$, $V$ denotes all possible subsets of $v$, $|V|$ denotes the size of $V$, and $\widehat{x_V}$ denotes the predicted value of the model for the data instance $x$ when all features are included in the feature subset $V$. $\hat{x}_{V \setminus v}$ denotes the prediction of a data instance $x$ given the exclusion of feature $v$ from the feature subset $V$. In this study, we calculated the SHAP value of each feature using the honeycomb graph to assess its importance in the classification decision.

### 2.7. Computational Environment

Spectral pre-processing and significant wavelength selection were performed using The Unscrambler X 0.4 (64-bit) and MATLAB 2021a, respectively. All modeling algorithms were developed via Python 3.9 and PyTorch 1.10.
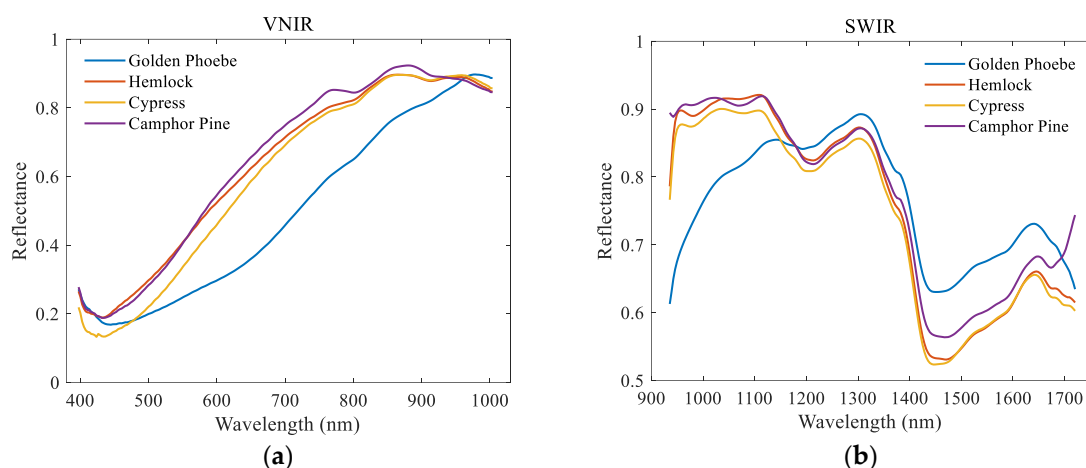
## 3. Results

### 3.1. Spectral Analysis

Figure 4 shows the average spectra of the different species of wood. Within the VNIR, absorption occurs near 450 nm, linked to pigments in the lignin and cellulose of wood. This absorption is due to the π-π* electron transitions of aromatic rings (C=C) and carbonyl groups (C=O) in the pigment molecules, leading to strong absorption of blue-violet light [36]. Additionally, there are differences in reflectance among various wood samples across the spectral range. For example, Golden Phoebe shows lower reflectance throughout the spectra, particularly in the 800–1000 nm range, likely due to its higher density and the increased light absorption caused by the C-H and C=C bonds in lignin and cellulose [37]. In contrast, Hemlock, Cypress, and Camphor Pine exhibit higher reflectance in the 600–900 nm range, which may be related to their looser fiber structure that enhances light scattering, combined with lower moisture and volatile component content, reducing light

absorption in this band. In the SWIR region, the reflectance also varies, with absorption peaks around 1100 nm, 1200 nm, and 1650 nm. The absorption peak near 1100 nm is linked to the vibrations of C-H bonds, particularly in the hydrocarbons of cellulose and hemicellulose. The peak around 1300 nm is primarily associated with the first vibrational overtone of O-H bonds, reflecting the moisture content and hydroxyl (O-H) content in cellulose. Lastly, the absorption peak near 1650 nm is related to the vibrational overtone frequencies of N-H and O-H bonds, with Golden Phoebe and Hemlock showing strong absorption, suggesting higher levels of lignin or nitrogen compounds in these samples [38]. It is worth noting that between 900 nm and 1200 nm, the reflectivity of Golden Phoebe is the lowest relative to the other three woods, which may be due to its high content of extracts (such as resins and aromatic compounds), which show strong absorption in this range.



**Figure 4.** Average reflectance spectra of different wood in (**a**) VNIR and (**b**) SWIR ranges.

*3.2. Principal Component Analysis*

PCA analyses were performed on the spectral data in the VNIR and SWIR ranges and the results are shown in Figure 5. In the VNIR range, the PCA results showed that the principal components PC−1, PC−2, and PC−3 explained 87.95%, 6.32%, and 3.83% of the data variance, respectively, with a cumulative contribution of 98.1%. As can be seen from the figure, Golden Phoebe showed a clear separation from the other wood samples on the PC-1 axis, indicating that its spectral characteristics in the VNIR spectral range were significantly different from those of the other woods. This separation may be related to the large color difference of Golden Phoebe. Although Hemlock, Cypress, and Camphor Pine are clustered to some extent on the PC−2 and PC-3 axes, their distribution on the PC-1 axis is more concentrated, suggesting that the spectral characteristics of these wood samples are more similar and less distinguishable in the VNIR range. In the SWIR range, the results of PCA analysis showed that the principal components PC−1, PC−2, and PC−3 explained 56.88%, 31.55%, and 5.60% of the data variance, respectively, with a cumulative contribution rate of 94.03%. There was a significant increase in the overlap between the wood samples in the SWIR spectral range, especially the samples of Hemlock, Cypress, and Camphor Pine had almost complete overlap in the PC−1 and PC−2 axes, which indicated that these wood samples had less differences in spectral characteristics in the SWIR range. This phenomenon may imply that the internal composition or structural characteristics of the wood samples in the SWIR spectral range show a high degree of similarity, resulting in the inability to distinguish between the wood samples in the PCA analysis. Therefore, PCA analysis alone may not be sufficient to accurately differentiate these wood samples, and a combination of other analytical methods or models is needed to improve the accuracy of the differentiation [39,40].
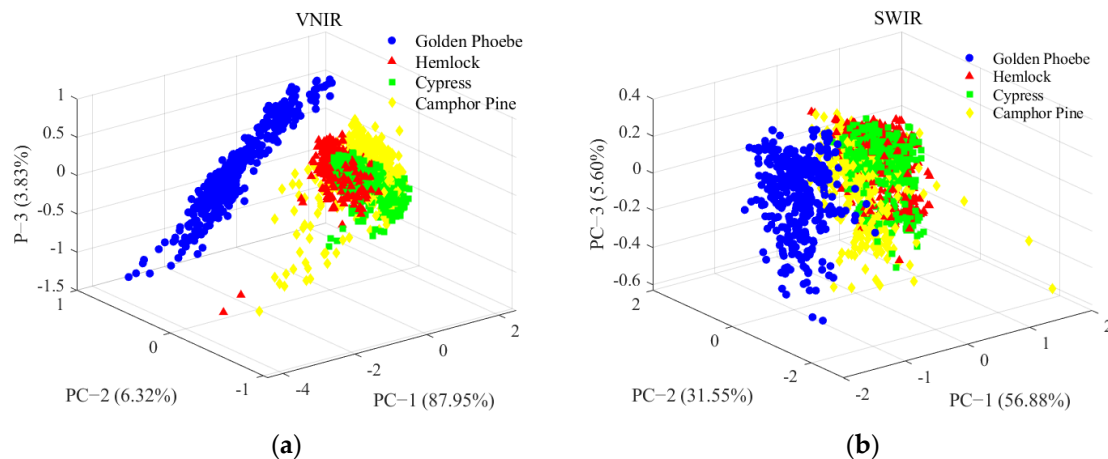
**Figure 5.** Results of PCA analysis of VNIR and SWIR datasets, (**a**) VNIR, (**b**) SWIR.

### 3.3. Validation of Species Classification Models

To better distinguish wood species, this study conducted validation analyses from two perspectives: full wavelength and feature wavelength. The testing utilized PLS-DA, XGBoost, CNN, and CNN-Transformer models. For the full wavelength analysis, various preprocessing methods, including SG, NOR, BL, SNV, and MSC were applied. For the feature wavelength analysis. Based on the most effective preprocessing method, feature extraction was performed using CARS, SPA, and RFE. Deep learning models are highly sensitive to initialization and random variations. To ensure the robustness of the results, this study conducted five repetitions of the experiment and presented the results by reporting the mean and standard deviation of the accuracy.

### 3.3.1. Full Wavelength Modeling

Table 2 presents the full-band modeling results. In the VNIR range, the NOR preprocessing method yielded the best performance for the PLS-DA and XGBoost models, with test set accuracies reaching 96.67% and 96.25%, respectively. In contrast, the CNN and CNN-Transformer models performed better on the unprocessed RAW data, achieving accuracies of 95.46 ± 1.65% and 98.21 ± 0.67%, respectively. This underscores the influence of preprocessing techniques on model outcomes. PLS-DA and XGBoost, as more conventional machine learning models, benefit from preprocessing methods like NOR, which reduce noise and emphasize linear patterns in the data, thereby enhancing these models' ability to identify relevant features [41]. On the other hand, CNN and CNN-Transformer, as deep learning architectures, possess a stronger inherent capacity for automatic feature extraction and can effectively manage raw or minimally preprocessed data, such as SNV, which preserves the complexity of the original spectral patterns. This highlights the role of these models' ability to extract features directly from the data while maintaining robustness against noise [42]. In the SWIR range, PLS-DA and XGBoost again performed well, with test set accuracies of 98.75% and 97.71%, respectively. However, the CNN and CNN-Transformer models demonstrated superior performance when employing the SNV preprocessing method, achieving accuracies of 98.96 ± 0.71% and 99.92 ± 0.19%, respectively. These findings suggest that selecting the appropriate combination of preprocessing methods and models is critical to optimizing classification performance [13]. In particular, the deep learning models, and CNN-Transformer in particular, are able to leverage the nuanced spectral information enhanced by SNV preprocessing, facilitating the extraction of complex patterns that are less accessible to traditional models. The superior performance of CNN-Transformer in the SWIR band further reflects its ability to capture detailed spectral features of wood species, especially those linked to chemical and molecular

structures, such as the absorption peaks associated with lignin and cellulose [43]. This ability to capture and differentiate subtle spectral variations contributes to the model's outstanding classification performance in the SWIR. Overall, recognition accuracy in the SWIR range is higher than in the VNIR range, especially in the case of the CNN-Transformer model.

**Table 2.** Modeling Results for Wood Species Classification Across VNIR and SWIR Using Different Preprocessing Methods and Models.

| Model | Method | VNIR-Accuracy/% | | SWIR-Accuracy/% | |
|---|---|---|---|---|---|
| | | **Train** | **Test** | **Train** | **Test** |
| **PLS-DA** | RAW | 92.32 | 95.21 | 95.71 | 95.42 |
| | SG | 91.34 | 95.21 | 95.18 | 95.21 |
| | NOR | 95.89 | **96.67** | 99.11 | 98.75 |
| | BL | 90.63 | 90.21 | 97.68 | **98.33** |
| | SNV | 94.02 | 96.25 | 92.59 | 92.08 |
| | MSC | 95.27 | 93.75 | 97.86 | 97.5 |
| XGBoost | RAW | 97.14 | 95.21 | 96.96 | 92.08 |
| | SG | 96.79 | 95.42 | 98.30 | 93.12 |
| | NOR | 99.38 | **96.25** | 100 | **97.71** |
| | BL | 97.68 | 94.38 | 99.55 | 96.46 |
| | SNV | 94.20 | 95.62 | 99.91 | 95.21 |
| | MSC | 95.54 | 81.04 | 99.20 | 85.83 |
| CNN | RAW | 99.01 ± 1.20 | 95.46 ± 1.65 | 98.29 ± 0.64 | 96.25 ± 0.98 |
| | SG | 97.86 ± 1.27 | 94.54 ± 0.50 | 99.04 ± 0.76 | 97.10 ± 0.41 |
| | NOR | 99.91 ± 0.11 | 94.50 ± 2.11 | 99.95 ± 0.08 | 97.71 ± 0.21 |
| | BL | 97.53 ± 1.90 | 94.63 ± 1.23 | 99.52 ± 0.83 | 97.75 ± 0.94 |
| | SNV | 98.18 ± 0.56 | 95.46 ± 0.18 | 99.43 ± 0.76 | **98.96 ± 0.71** |
| | MSC | 98.67 ± 0.79 | 93.50 ± 1.05 | 99.27 ± 0.96 | 94.16 ± 0.53 |
| CNN-Transformer | RAW | 97.09 ± 0.63 | **98.21 ± 0.67** | 99.61 ± 0.46 | 98.84 ± 0.52 |
| | SG | 96.61 ± 1.09 | 97.46 ± 0.63 | 95.77 ± 1.77 | 95.42 ± 2.88 |
| | NOR | 95.52 ± 1.02 | 94.83 ± 1.28 | 98.07 ± 0.67 | 97.92 ± 1.47 |
| | BL | 97.70 ± 0.27 | 97.04 ± 0.47 | 99.32 ± 0.53 | 99.08 ± 0.60 |
| | SNV | 97.32 ± 0.83 | 96.79 ± 0.38 | 99.98 ± 0.04 | **99.92 ± 0.19** |
| | MSC | 96.14 ± 1.42 | 95.46 ± 0.37 | 93.05 ± 1.41 | 89.07 ± 0.85 |

Note: The bolded part represents the best modeling result

In order to gain insight into the classification performance of the model, one of the five experiments with the best results was selected and a confusion matrix was generated, as shown in Figure 6. In the VNIR range, none of the four models misclassified Golden Phoebe, likely due to its distinct color and spectral characteristics. The VNIR band's sensitivity to surface reflectance may enhance the models' ability to distinguish Golden Phoebe from other species. However, Hemlock, Cypress, and Camphor Pine were misclassified to varying degrees in the PLS-DA, XGBoost, and CNN models, likely because of their similar surface textures and microscopic spectral absorption properties in the VNIR band. Although CNN-Transformer demonstrated better overall performance, misclassification still occurred with Hemlock and Cypress, which could be attributed to the similarity in their texture and absorption features within the VNIR band. In contrast, in the SWIR range, the PLS-DA and XGBoost models exhibited misclassification due to overlapping spectral features in the 1200–1600 nm range. These traditional models rely on linear feature extraction, which makes distinguishing small differences challenging when spectral features overlap significantly. Similarly, the CNN also showed misclassification

when distinguishing Hemlock and Cypress, likely because their lignin and hemicellulose content are highly similar in the SWIR range. The spectral absorption characteristics of these two species overlap, particularly in the critical 1200–1600 nm range, making it difficult for the CNN to extract distinguishing features. However, the CNN-Transformer model was able to capture these subtle differences more effectively. Combining convolutional layers and self-attention mechanisms, the CNN-Transformer is better suited to detect minor spectral variations, which explains its superior classification accuracy and lack of misclassification in the SWIR range.
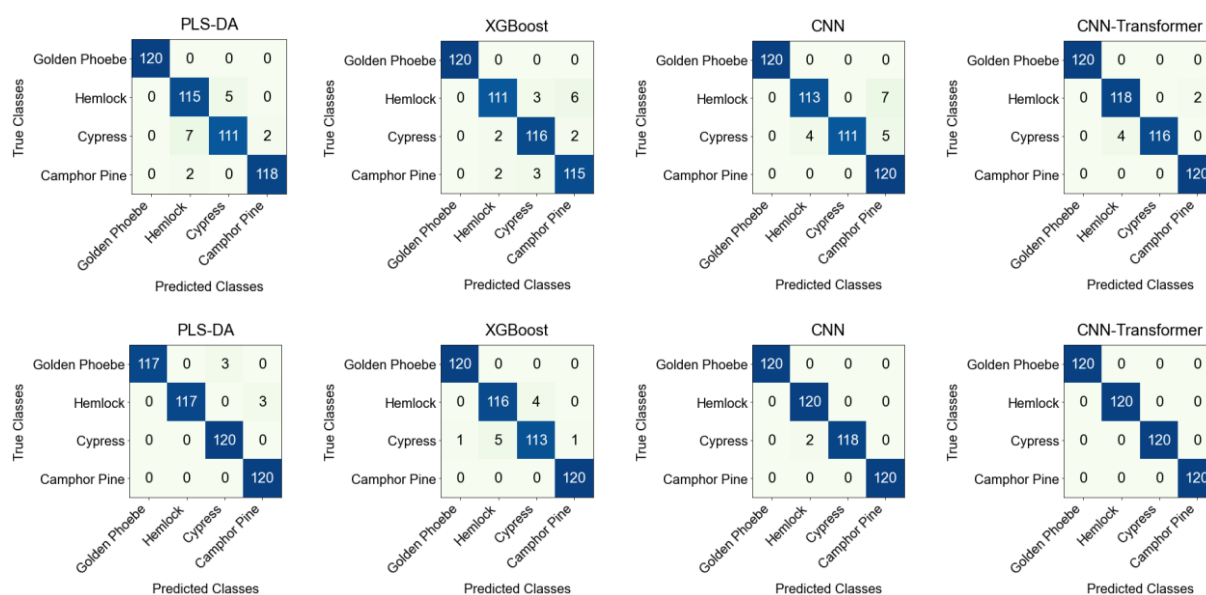


**Figure 6.** Confusion matrix of the four model results, with the first row representing VNIR and the second row SWIR.

In previous analyses, the classification performance of each model across different spectral bands was evaluated using confusion matrices. To further validate these findings, the Precision, Recall, and F1-Score for each model under the VNIR and SWIR were calculated, as shown in Table 3. Notably, discrepancies between Precision and Recall were observed in some model evaluations, revealing distinctive performance characteristics in various scenarios. For example, in the case of the CNN model's recognition of Camphor Pine in the VNIR, while Recall reached 100%, indicating that the model identified all Camphor Pine samples without omissions, Precision was only 90.91%. This suggests that nearly 10% of the samples classified as Camphor Pine were misclassified from other categories, highlighting the model's limitations in distinguishing similar species and introducing a risk of misclassification in its attempt to maximize identification. Similarly, the asymmetry between Precision (95.08%) and Recall (96.67%) in XGBoost's recognition of Cypress in the VNIR band warrants attention. This imbalance suggests that the model is somewhat conservative in correctly identifying the category, resulting in a degree of missed detection. While the overall accuracy remains high, such biases in identifying specific species may impact the model's performance in practical applications. In summary, the differences between Precision and Recall underscore the challenges models face in distinguishing between similar wood species. High overall accuracy may not fully capture the model's true performance, particularly in scenarios where spectral features overlap, leading to potential instability in classification.

**Table 3.** Precision, Recall, and F1-Score of Models for Wood Species Identification under VNIR and SWIR.

| Range | Model | Indicator/% | Species | | | |
|---|---|---|---|---|---|---|
| | | | Golden Phoebe | Hemlock | Cypress | Camphor Pine |
| VNIR | PLS-DA | Precision | 100 | 92.74 | 95.69 | 98.33 |
| | | Recall | 100 | 95.83 | 92.50 | 98.33 |
| | | F1-Score | 100 | 94.26 | 94.07 | 98.33 |
| | XGBoost | Precision | 100 | 96.52 | 95.08 | 93.50 |
| | | Recall | 100 | 92.50 | 96.67 | 95.83 |
| | | F1-Score | 100 | 94.47 | 95.87 | 94.65 |
| | CNN | Precision | 100 | 96.58 | 100 | 90.91 |
| | | Recall | 100 | 94.17 | 92.50 | 100 |
| | | F1-Score | 100 | 95.36 | 96.10 | 95.24 |
| | CNN-Transformer | Precision | 100 | 96.72 | 98.31 | 100 |
| | | Recall | 100 | 98.33 | 96.67 | 100 |
| | | F1-Score | 100 | 97.52 | 97.48 | 100 |
| SWIR | PLS-DA | Precision | 100 | 100 | 97.56 | 97.56 |
| | | Recall | 97.50 | 97.50 | 100 | 100 |
| | | F1-Score | 98.73 | 98.73 | 98.77 | 98.77 |
| | XGBoost | Precision | 99.17 | 95.87 | 96.58 | 99.17 |
| | | Recall | 100 | 96.67 | 94.17 | 100 |
| | | F1-Score | 99.59 | 96.27 | 95.36 | 99.59 |
| | CNN | Precision | 100 | 98.36 | 100 | 100 |
| | | Recall | 100 | 100 | 98.33 | 100 |
| | | F1-Score | 100 | 99.17 | 99.16 | 100 |
| | CNN-Transformer | Precision | 100 | 100 | 100 | 100 |
| | | Recall | 100 | 100 | 100 | 100 |
| | | F1-Score | 100 | 100 | 100 | 100 |

In addition, Table 4 shows that different models have obvious differences in parameters, model size and FLOPs. PLS-DA has the least number of parameters, only 2,280, and a total model size of 0.01 MB, but its FLOPs is 5.11M, with a high computational complexity, which is more suitable for small-scale tasks, but has limitations in efficient deployment. XGBoost has 204,700 parameters, a model size of 0.78 MB, and 1.12M FLOPs, achieving a good balance between computational efficiency and model complexity. The number of parameters and complexity of deep learning models are relatively higher. The number of parameters of CNN reaches 1,739,092, the total model size is 6.64 MB, and the FLOPs is 1.99M, showing its high demand for storage space and computing resources. In comparison, CNN-Transformer has 1,286,228 parameters, a model size of 4.93 MB, and FLOPs reduced to 1.29M, which reduces computational complexity and storage requirements while maintaining high performance. This efficiency improvement is due to the advantages of the Transformer structure in capturing key features and global dependencies. In general, CNN-Transformer strikes a good balance between efficiency and performance, and is very suitable for tasks such as wood species identification that require high accuracy and efficiency.

**Table 4.** Comparison of Computational Metrics Across Different Models.

| Model | Total Parameters | Model Size (MB) | FLOPs (M) |
|---|---|---|---|
| PLS-DA | 2,280 | 0.01 | 5.11 |
| XGBoost | 204,700 | 0.78 | 1.12 |
| CNN | 1,739,092 | 6.64 | 1.99 |
| CNN-Transformer | 1,286,228 | 4.93 | 1.29 |

3.3.2. Feature Extraction

The identification of wood species fundamentally relies on the classification ability of the model, and an in-depth analysis of classification properties is crucial for ensuring identification accuracy. However, covariance and redundancy among different wavelengths inevitably increase computational complexity, complicating model stability [44]. In this study, feature extraction from VNIR and SWIR data were conducted using CARS, SPA, and RFE methods, with the distribution of feature wavelengths illustrated in Figure 7. In the VNIR, 15, 11, and 24 feature wavelengths were extracted, primarily concentrated in the 400–600 nm and 800–1000 nm ranges, which are mainly associated with wood pigmentation, moisture content, and the surface properties of the fiber structure. For the SWIR, 21, 13, and 21 feature wavelengths were extracted, mainly around 1100 nm, 1300 nm, and 1650 nm, corresponding primarily to lignin, cellulose, and hemicellulose in wood. Additionally, the wavelengths extracted by the three feature extraction methods showed considerable overlap, and these extracted feature wavelengths will be used for further modeling.



(**a**)                                                                 (**b**)
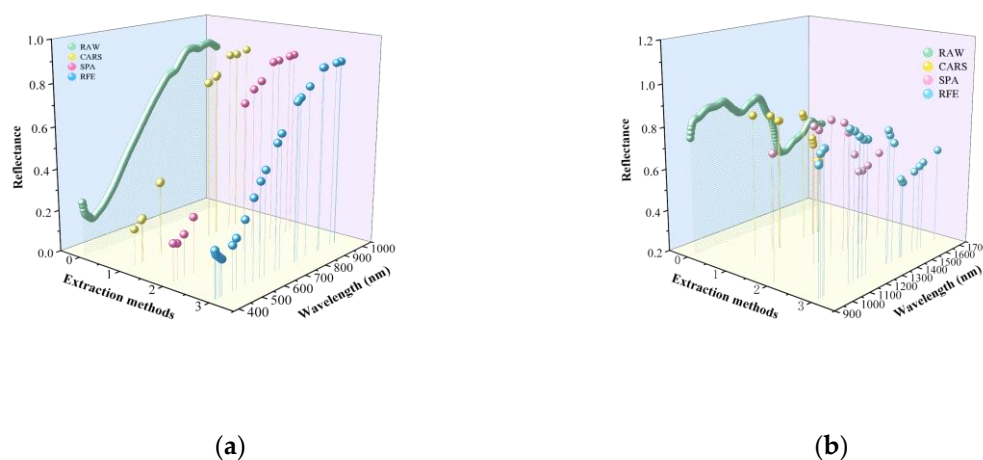
**Figure 7.** Specific locations of feature wavelengths extracted by CARS, SPA, and RFE based on the (**a**) VNIR data and (**b**) the SWIR data.

3.3.3. Feature Wavelength Modeling

Table 5 demonstrates the statistical results of each model after combining CARS, SPA, and RFE feature extraction methods. In the VNIR band, the performance of the PLS-DA, CNN, and CNN-Transformer models decreases compared to the full band modeling, and only the XGBoost model improves its performance with an accuracy of 97.92%. In the SWIR band, the results are different. While the performance of the PLS-DA model improved to 99.58%, the performance of the XGBoost, CNN, and CNN-Transformer models all decreased. These results indicate that although feature extraction methods have advantages in simplifying the model and reducing the computational complexity, they may also lead to the loss of some important bands, which affects the stability of the model and the classification accuracy [45]. The selected bands in the feature extraction process may

not adequately represent the spectral information of the full band, and this instability is particularly prominent when dealing with classification tasks with complex spectral features. Specifically, XGBoost shows stronger robustness in the VNIR band, which may be related to its better adaptation to feature selection; while in the SWIR band, the improved performance of PLS-DA may stem from its better ability to capture linear relationships. However, for complex nonlinear models such as CNN and CNN-Transformer, the loss of information during feature extraction may negatively affect the model performance. In order to further explore the decision-making mechanism of the model, the results of the visualization of the model are discussed in detail in the next section.

**Table 5.** Statistical results of the model combined with different wavelength selection methods.

| Model | Method | VNIR-Accuracy/% | | SWIR-Accuracy/% | |
|---|---|---|---|---|---|
| | | **Train** | **Test** | **Train** | **Test** |
| PLS-DA | CARS | 92.23 | 93.54 | 98.21 | 99.58 |
| | SPA | 92.50 | 94.38 | 99.64 | 99.58 |
| | RFE | 94.02 | 93.96 | 92.59 | 92.71 |
| XGBoost | CARS | 98.93 | 96.25 | 100 | 96.67 |
| | SPA | 99.64 | 97.50 | 100 | 97.29 |
| | RFE | 99.20 | 97.92 | 100 | 93.54 |
| CNN | CARS | 95.61 ± 0.52 | 94.29 ± 0.72 | 92.41 ± 1.15 | 91.04 ± 0.42 |
| | SPA | 86.17 ± 0.34 | 91.75 ± 1.39 | 98.23 ± 0.19 | 97.42 ± 0.24 |
| | RFE | 98.13 ± 0.97 | 94.58 ± 0.33 | 93.79 ± 0.46 | 91.42 ± 1.92 |
| CNN-Transformer | CARS | 94.95 ± 1.29 | 93.14 ± 1.52 | 92.86 ± 1.18 | 90.38 ± 2.76 |
| | SPA | 93.73 ± 0.95 | 90.54 ± 1.86 | 93.05 ± 0.77 | 90.71 ± 3.01 |
| | RFE | 92.22 ± 0.67 | 90.75 ± 1.90 | 94.48 ± 0.70 | 91.67 ± 1.05 |

*3.4. Explanation of Model*

SHAP was employed to interpret the model outputs by analyzing the contribution of each input feature to the predictions. SHAP values highlight the importance of individual features in wood classification predictions [46]. Figure 8 shows the distribution of SHAP values for the top 20 wavelengths in the CNN-Transformer model, while the results for the other three models are presented in Figure S1. Although the full spectral data contains 224 wavelengths, only the 20 most influential wavelengths are displayed for clarity. The use of SHAP values reveals the critical role these wavelengths play in the model's predictions.

Figure 8a shows that, in the VNIR range, several wavelengths in the range of 640–660 nm (e.g., 647.58 nm and 652.99 nm), 700–720 nm, and 790–800 nm (e.g., 718.15 nm and 794.84 nm) significantly impact the model's output. These wavelengths correspond to the absorption characteristics of key wood components, such as lignin and cellulose. For instance, 647.58 nm and 652.99 nm are closely related to the absorption of pigments and structural components in the wood, which are indicative of specific wood types with higher lignin content [36]. The broad SHAP value distributions for these wavelengths suggest that they play a critical role in the model's decision-making process by capturing variations in wood density and composition. Additionally, wavelengths in the range of 590–600 nm and 830–840 nm, although less influential, reflect important features such as the fiber structure and surface reflectivity of the wood, which contribute to classification accuracy. Figure 8b highlights the significance of multiple wavelengths in the SWIR range. Specifically, wavelengths in the range of 1040–1070 nm (e.g., 1049.95 nm), 1280–1320 nm, and 1350–1370 nm (e.g., 1294.53 nm and 1364.92 nm) substantially contribute to the model's predictions. These ranges correspond to the molecular vibrations of C-H and O-H bonds, primarily in the lignin and hemicellulose structures of the wood. The SHAP

value distributions for these wavelengths emphasize their importance in identifying subtle differences in wood species based on their chemical composition. Furthermore, wavelengths in the range of 960–980 nm and 1700–1720 nm, although less pronounced, likely correspond to specific moisture content and nitrogen-related compounds in the wood, which also influence the model's classification decisions.
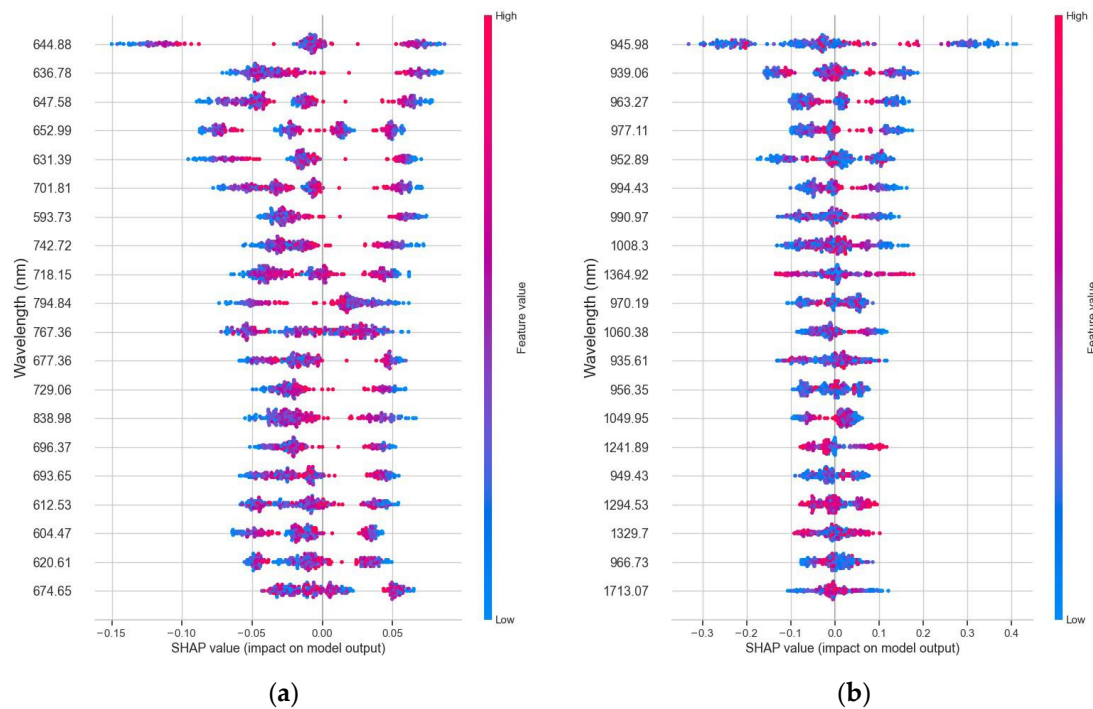


**Figure 8.** SHAP feature importance plot for predicting wood species. (**a**) VNIR. (**b**) SWIR.

A comparison of the PLS-DA and XGBoost models in Figure S1 shows that these traditional models also identify key wavelengths. For example, in the VNIR band, the PLS-DA model assigns higher SHAP values to wavelengths within the 440–480 nm range, which are associated with the absorption of pigments and surface-level features of wood. The XGBoost model displays a similar trend in this range, indicating that both models rely on specific, narrow wavelength regions for classification. However, compared to CNN and CNN-Transformer models, PLS-DA, and XGBoost show a more concentrated pattern of wavelength selection. This focus on a limited number of critical wavelengths suggests that traditional models may depend more heavily on specific spectral bands, which explains their improved performance after feature extraction. By focusing on key wavelengths, these models reduce noise and improve accuracy, albeit at the cost of losing broader spectral information. In contrast, deep learning models, such as CNN and CNN-Transformer, leverage more distributed and diverse wavelength ranges. In the SWIR band, the CNN-Transformer model not only identifies significant wavelengths in the 1280–1370 nm range but also captures relevant spectral features across a broader range (e.g., 1000–1400 nm). This contrasts sharply with the PLS-DA and XGBoost models, where high SHAP values are concentrated within relatively narrow ranges, primarily around 1430–1480 nm and 1660–1720 nm. The wider distribution of influential wavelengths in the CNN-Transformer model highlights its ability to capture more nuanced spectral variations that are linked to the physicochemical properties of wood [47]. This broader capability allows deep learning models to extract spectral information that is often missed by traditional models, which enhances both classification accuracy and robustness.

While Figure 8 only highlights the top 20 influential wavelengths, it is important to note that the CNN-Transformer model analyzes and leverages spectral data across a much

wider range. This broader capability allows the model to capture subtle yet critical spectral variations linked to wood's molecular structure. In the VNIR range, wavelengths such as 647.58 nm and 652.99 nm may correspond to the pigmentation and surface texture of the wood, which is essential for distinguishing between visually similar species [48]. In contrast, in the SWIR range, wavelengths like 1049.95 nm and 1364.92 nm are likely connected to deeper chemical components, such as lignin and cellulose, that contribute to the structural integrity of the wood.

## 4. Discussion

This study validates the potential of hyperspectral imaging combined with deep learning models in wood species classification, especially in the SWIR band, where the CNN-Transformer model demonstrates excellent performance. This result is consistent with recent studies on the application of deep learning in spectral data classification, showing that deep learning models, especially the Transformer architecture combined with the self-attention mechanism, can effectively extract global features from high-dimensional spectral data and overcome the limitations of traditional machine learning methods in high-dimensional data [49]. Compared with traditional machine learning models such as XGBoost and PLS-DA, CNN-Transformer not only extracts local features through the convolutional layer, but also captures the complex relationship between different wavelengths through the self-attention mechanism, which demonstrates a clear advantage, especially in the region with strong spectral overlap. For example, in the SWIR band, the classification accuracy of the traditional model decreases significantly when confronted with similar species such as Hemlock, Cypress, and Camphor Pine, while the CNN-Transformer is able to accurately distinguish these species through finer feature extraction and global modeling, a phenomenon which further demonstrates the deep learning model's powerful ability in processing complex spectral data. From a computational perspective, the CNN-Transformer model strikes an effective balance between performance and efficiency. As shown in Table 4, it reduces the total number of parameters and computational complexity (FLOPs) compared to CNN, while maintaining a compact model size. This makes CNN-Transformer suitable for resource-constrained scenarios, such as embedded systems or portable devices used in the wood industry. These computational advantages further enhance its applicability in real-world tasks where both accuracy and efficiency are paramount.

However, despite the excellent performance of deep learning models, the role of feature selection in wood species classification should not be ignored. Feature selection methods such as CARS, SPA, and RFE can effectively reduce the data dimensionality and alleviate the computational burden of the model, especially in traditional models, and feature selection effectively improves the stability and performance of the model [50]. In this study, although the feature extraction method reduces the dimensionality of the data, the feature selection oversimplifies the spectral data for deep learning models such as CNN-Transformer, resulting in a certain degree of information loss, especially in the SWIR bands, where the accuracy of CNN-Transformer decreases from 99.92 ± 0.19% to 91.67 ± 1.05%. This result is in line with the study by [51], who noted that while feature selection can improve performance in traditional models, over-simplification of data for deep learning models may affect their ability to capture complex spectral features. In contrast, XGBoost's performance improves after feature selection, indicating that it is more adaptable to simplified data, especially in the VNIR band, with an accuracy of 97.92%. This phenomenon further suggests that traditional models tend to function more consistently in the face of data simplification, whereas deep learning models rely more on data completeness and diversity.

In addition, the interpretability of the model was explored in this study through SHAP value analysis. The CNN-Transformer model performed outstandingly in the SHAP value analysis and was able to identify key wavelengths associated with the chemical composition of wood (e.g., lignin, cellulose), especially in the SWIR bands, and the spectral features corresponding to these wavelengths had a significant impact on the classification results. Compared to conventional models, CNN-Transformer integrates features at multiple wavelengths through a self-attentive mechanism, which enables it to capture the molecular structure and compositional features of wood in a more comprehensive way. However, although the SHAP values provide some interpretability to the model, the complexity of the CNN-Transformer still makes its decision-making process not fully transparent, especially in the high-dimensional environment of spectral data, and a more intuitive understanding of the internal mechanisms of the model remains a major challenge for deep learning in practical applications. In contrast, traditional models such as PLS-DA and XGBoost rely on more concentrated wavelength intervals after feature selection, and thus, have a high degree of interpretability, and are able to make quick judgments through features in a single band, which is why they are stable and effective in the face of simpler or low-dimensional data.

This study also suffers from the fact that the wood source is relatively homogeneous, which limits the generalization ability of the model. Although the results show that CNN-Transformer has excellent classification performance on this sample set, the performance of the model on wood samples from other sources has not been validated. Interregional ecological and environmental conditions may significantly affect the spectral properties of wood, which in turn challenges the robustness of the model. In addition, aging or degradation of wood may alter its spectral properties, and changes in moisture content, especially in the SWIR band, can significantly affect physical and spectral properties. Future studies need to incorporate wood samples from different regions, ages, and moisture contents to systematically evaluate the generalization ability and robustness of the model, and explore the enhancement of model stability through domain adaptation techniques or improved feature extraction methods. This will help improve the applicability and reliability of the model in diverse application scenarios.

## 5. Conclusions

This study demonstrated the effectiveness of combining hyperspectral imaging (HSI) technology with artificial intelligence algorithms for wood species identification. Among the tested methods, the CNN-Transformer model achieved superior classification performance on SWIR data, with an accuracy of $99.92 \pm 0.19\%$, highlighting its ability to capture complex spectral features. SHAP analysis further emphasized the importance of specific spectral regions, providing insights into the model's interpretability. Comparatively, traditional models such as PLS-DA and XGBoost exhibited limitations in handling intricate spectral data despite improvements in feature extraction. These findings establish the combination of SWIR and CNN-Transformer as a reliable approach for accurate wood species classification, providing a solid foundation for the development of automated and intelligent wood processing systems.

**Author Contributions:** Y.Q.: conceptualization, methodology, investigation, funding acquisition, writing—original draft, Y.Z.: investigation, formal analysis, methodology. S.T.: investigation,

validation, software, writing—original draft, Z.Z.: data curation, writing—review and editing. All authors have read and agreed to the published version of the manuscript.

**Data Availability Statement:** The authors do not have permission to share data.

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

# References

1. Shrestha, S.; Kognou, A.L.M.; Zhang, J.; Qin, W. Different Facets of Lignocellulosic Biomass Including Pectin and Its Perspectives. *Waste Biomass Valoriz.* **2021**, *12*, 4805–4823.

2. Zhou, Z.; Rahimi, S.; Avramidis, S. On-Line Species Identification of Green Hem-Fir Timber Mix Based on near Infrared Spectroscopy and Chemometrics. *Eur. J. Wood Wood Prod.* **2020**, *78*, 151–160. https://doi.org/10.1007/s00107-019-01479-8.

3. Hu, J.L.; Ci, X.Q.; Liu, Z.F.; Dormontt, E.E.; Conran, J.G.; Lowe, A.J.; Li, J. Assessing Candidate DNA Barcodes for Chinese and Internationally Traded Timber Species. *Mol. Ecol. Resour.* **2022**, *22*, 1478–1492. https://doi.org/10.1111/1755-0998.13546.

4. Ma, T.; Inagaki, T.; Ban, M.; Tsuchikawa, S. Rapid Identification of Wood Species by Near-Infrared Spatially Resolved Spectroscopy (NIR-SRS) Based on Hyperspectral Imaging (HSI). *Holzforschung* **2019**, *73*, 323–330. https://doi.org/10.1515/hf-2018-0128.

5. Tsuchikawa, S.; Kobori, H. A Review of Recent Application of near Infrared Spectroscopy to Wood Science and Technology. *J. Wood Sci.* **2015**, *61*, 213–220. https://doi.org/10.1007/s10086-015-1467-x.

6. Schimleck, L.; Ma, T.; Inagaki, T.; Tsuchikawa, S. Review of near Infrared Hyperspectral Imaging Applications Related to Wood and Wood Products. *Appl. Spectrosc. Rev.* **2023**, *58*, 585–609.

7. Mancini, M.; Taavitsainen, V.M.; Rinnan, Å. Comparison of Classification Methods Performance for Defining the Best Reuse of Waste Wood Material Using NIR Spectroscopy. *Waste Manag.* **2024**, *178*, 321–330. https://doi.org/10.1016/j.wasman.2024.02.033.

8. Goi, A.; Hocquette, J.F.; Pellattiero, E.; De Marchi, M. Handheld Near-Infrared Spectrometer Allows on-Line Prediction of Beef Quality Traits. *Meat Sci.* **2022**, *184*, 108694. https://doi.org/10.1016/j.meatsci.2021.108694.

9. Ren, Z.; Zhang, Z.; Wei, J.; Dong, B.; Lee, C. Wavelength-Multiplexed Hook Nanoantennas for Machine Learning Enabled Mid-Infrared Spectroscopy. *Nat. Commun.* **2022**, *13*, 3859. https://doi.org/10.1038/s41467-022-31520-z.

10. Agarwal, U.P.; Ralph, S.A.; Reiner, R.S.; Baez, C. Probing Crystallinity of Never-Dried Wood Cellulose with Raman Spectroscopy. *Cellulose* **2016**, *23*, 125–144. https://doi.org/10.1007/s10570-015-0788-7.

11. Ho, T.X.; Schimleck, L.R.; Dahlen, J.; Sinha, A. Utilization of Genetic Algorithms to Optimize Loblolly Pine Wood Property Models Based on NIR Spectra and SilviScan Data. *Wood Sci. Technol.* **2022**, *56*, 1419–1437. https://doi.org/10.1007/s00226-022-01403-z.

12. Lima, M.D.R.; Ramalho, F.M.G.; Trugilho, P.F.; Bufalino, L.; Dias Júnior, A.F.; Protásio, T.d.P.; Hein, P.R.G. Classifying Waste Wood from Amazonian Species by Near-Infrared Spectroscopy (NIRS) to Improve Charcoal Production. *Renew. Energy* **2022**, *193*, 584–594. https://doi.org/10.1016/j.renene.2022.05.048.

13. Li, P.; Tang, S.; Chen, S.; Tian, X.; Zhong, N. Hyperspectral Imaging Combined with Convolutional Neural Network for Accurately Detecting Adulteration in Atlantic Salmon. *Food Control* **2023**, *147*, 109573. https://doi.org/10.1016/j.foodcont.2022.109573.

14. Tang, S.; Zhang, L.; Tian, X.; Zheng, M.; Su, Z.; Zhong, N. Rapid Non-Destructive Evaluation of Texture Properties Changes in Crispy Tilapia during Crispiness Using Hyperspectral Imaging and Data Fusion. *Food Control* **2024**, *162*, 110446. https://doi.org/10.1016/j.foodcont.2024.110446.

15. Feng, L.; Wu, B.; He, Y.; Zhang, C. Hyperspectral Imaging Combined With Deep Transfer Learning for Rice Disease Detection. *Front. Plant Sci.* **2021**, *12*, 693521. https://doi.org/10.3389/fpls.2021.693521.

16. Feng, Y.; Lv, Y.; Dong, F.; Chen, Y.; Li, H.; Rodas-González, A.; Wang, S. Combining Vis-NIR and NIR Hyperspectral Imaging Techniques with a Data Fusion Strategy for Prediction of Norfloxacin Residues in Mutton. *Spectrochim. Acta A Mol. Biomol. Spectrosc.* **2024**, *322*, 124844. https://doi.org/10.1016/j.saa.2024.124844.

17. Yadav, C.S.; Pradhan, M.K.; Gangadharan, S.M.P.; Chaudhary, J.K.; Singh, J.; Khan, A.A.; Haq, M.A.; Alhussen, A.; Wechtaisong, C.; Imran, H.; et al. Multi-Class Pixel Certainty Active Learning Model for Classification of Land Cover Classes Using Hyperspectral Imagery. *Electronics* **2022**, *11*, 2799. https://doi.org/10.3390/electronics11172799.

18. Ma, P.; Jia, X.; Xu, W.; He, Y.; Tarwa, K.; Alharbi, M.O.; Wei, C.I.; Wang, Q. Enhancing Salmon Freshness Monitoring with Sol-Gel Cellulose Nanocrystal Colorimetric Paper Sensors and Deep Learning Methods. *Food Biosci.* **2023**, *56*, 103313. https://doi.org/10.1016/j.fbio.2023.103313.

19. Tetko, I.V.; Karpov, P.; Van Deursen, R.; Godin, G. State-of-the-Art Augmented NLP Transformer Models for Direct and Single-Step Retrosynthesis. *Nat. Commun.* **2020**, *11*, 5575. https://doi.org/10.1038/s41467-020-19266-y.

20. Zheng, Z.; Ge, Z.; Tian, Z.; Yang, X.; Zhou, Y. WoodGLNet: A Multi-Scale Network Integrating Global and Local Information for Real-Time Classification of Wood Images. *J. Real. Time Image Process* **2024**, *21*, 147. https://doi.org/10.1007/s11554-024-01521-w.

21. Zhao, N.; Duan, Z.; Li, Q.; Guo, K.; Zhang, Z.; Liu, B. A Cable Insulation Defect Classification Method Based on CNN-Transformer. *Front. Phys.* **2024**, *12*, 1432527. https://doi.org/10.3389/fphy.2024.1432527.

22. Zhang, X.; He, C.; Lu, Y.; Chen, B.; Zhu, L.; Zhang, L. Fault Diagnosis for Small Samples Based on Attention Mechanism. *Measurement* **2022**, *187*, 110242. https://doi.org/10.1016/j.measurement.2021.110242.

23. Li, Z. Extracting Spatial Effects from Machine Learning Model Using Local Interpretation Method: An Example of SHAP and XGBoost. *Comput. Environ. Urban Syst.* **2022**, *96*, 101845. https://doi.org/10.1016/j.compenvurbsys.2022.101845.

24. Chen, W.; Chen, H.; Feng, Q.; Mo, L.; Hong, S. A Hybrid Optimization Method for Sample Partitioning in Near-Infrared Analysis. *Spectrochim. Acta A Mol. Biomol. Spectrosc.* **2021**, *248*, 119182. https://doi.org/10.1016/j.saa.2020.119182.

25. Pan, X.; Li, K.; Chen, Z.; Yang, Z. Identifying Wood Based on Near-Infrared Spectra and Four Gray-Level Co-Occurrence Matrix Texture Features. *Forests* **2021**, *12*, 1527. https://doi.org/10.3390/f12111527.

26. Tuncer, F.D.; Dogu, D.; Akdeniz, E. Efficiency of Preprocessing Methods for Discrimination of Anatomically Similar Pine Species by NIR Spectroscopy. *Wood Mater. Sci. Eng.* **2023**, *18*, 212–221. https://doi.org/10.1080/17480272.2021.2012821.

27. Li, X.; Cai, M.; Li, M.; Wei, X.; Liu, Z.; Wang, J.; Jia, K.; Han, Y. Combining Vis-NIR and NIR Hyperspectral Imaging Techniques with a Data Fusion Strategy for the Rapid Qualitative Evaluation of Multiple Qualities in Chicken. *Food Control* **2023**, *145*, 109416. https://doi.org/10.1016/j.foodcont.2022.109416.

28. Li, Y.; Wang, G.; Guo, G.; Li, Y.; Via, B.K.; Pei, Z. Spectral Pre-Processing and Multivariate Calibration Methods for the Prediction of Wood Density in Chinese White Poplar by Visible and Near Infrared Spectroscopy. *Forests* **2022**, *13*, 62. https://doi.org/10.3390/f13010062.

29. Chen, J.; Yu, H.; Jiang, D.; Zhang, Y.; Wang, K. A Novel NIRS Modelling Method with OPLS-SPA and MIX-PLS for Timber Evaluation. *J. For. Res.* **2022**, *33*, 369–376. https://doi.org/10.1007/s11676-021-01314-y.

30. Ramezan, C.A. Transferability of Recursive Feature Elimination (RFE)-Derived Feature Sets for Support Vector Machine Land Cover Classification. *Remote Sens.* **2022**, *14*, 6218. https://doi.org/10.3390/rs14246218.

31. Leandro, J.G.R.; Gonzaga, F.B.; Latorraca, J.V.d.F. Discrimination of Wood Species Using Laser-Induced Breakdown Spectroscopy and near-Infrared Reflectance Spectroscopy. *Wood Sci. Technol.* **2019**, *53*, 1079–1091. https://doi.org/10.1007/s00226-019-01119-7.

32. Wang, S.; Liu, S.; Zhang, J.; Che, X.; Yuan, Y.; Wang, Z.; Kong, D. A New Method of Diesel Fuel Brands Identification: SMOTE Oversampling Combined with XGBoost Ensemble Learning. *Fuel* **2020**, *282*, 118848. https://doi.org/10.1016/j.fuel.2020.118848.

33. He, X.; Chen, Y.; Lin, Z. Spatial–Spectral Transformer for Hyperspectral Image Classification. *Remote Sens.* **2021**, *13*, 498. https://doi.org/10.3390/rs13030498.

34. Ding, K.; Lu, T.; Fu, W.; Li, S.; Ma, F. Global-Local Transformer Network for HSI and LiDAR Data Joint Classification. *IEEE Trans. Geosci. Remote Sens.* **2022**, *60*, 5541213. https://doi.org/10.1109/TGRS.2022.3216319.

35. Ahmed, T.; Wijewardane, N.K.; Lu, Y.; Jones, D.S.; Kudenov, M.; Williams, C.; Villordon, A.; Kamruzzaman, M. Advancing Sweetpotato Quality Assessment with Hyperspectral Imaging and Explainable Artificial Intelligence. *Comput. Electron. Agric.* **2024**, *220*, 108855. https://doi.org/10.1016/j.compag.2024.108855.

36. Muhammad, N.A.; Isnaeni; Tahir, D. Optical Properties of Wood by Laser Spectroscopy. In *Journal of Physics: Conference Series*; Institute of Physics Publishing: Bristol, UK, 2019; Volume 1341.

37. Qi, W.; Xiong, Z.; Tang, H.; Lu, D.; Chen, B. Compact Near-Infrared Spectrometer for Quantitative Determination of Wood Composition. *J. Appl. Spectrosc.* **2021**, *88*, 461–467. https://doi.org/10.1007/s10812-021-01194-4.

38. Peng, H.; Salmén, L.; Stevanic, J.S.; Lu, J. Structural Organization of the Cell Wall Polymers in Compression Wood as Revealed by FTIR Microspectroscopy. *Planta* **2019**, *250*, 163–171. https://doi.org/10.1007/s00425-019-03158-7.

39.	Sharma, V.; Yadav, J.; Kumar, R.; Tesarova, D.; Ekielski, A.; Mishra, P.K. On the Rapid and Non-Destructive Approach for Wood Identification Using ATR-FTIR Spectroscopy and Chemometric Methods. *Vib. Spectrosc.* **2020**, *110*, 103097. https://doi.org/10.1016/j.vibspec.2020.103097.

40.	Park, S.Y.; Kim, J.C.; Yeon, S.; Yang, S.Y.; Yeo, H.; Choi, I.G. Rapid Prediction of the Chemical Information of Wood Powder from Softwood Species Using Near-Infrared Spectroscopy. *Bioresources* **2018**, *13*, 2440–2451. https://doi.org/10.15376/biores.13.2.2440-2451.

41.	Tsakiridis, N.L.; Keramaris, K.D.; Theocharis, J.B.; Zalidis, G.C. Simultaneous Prediction of Soil Properties from VNIR-SWIR Spectra Using a Localized Multi-Channel 1-D Convolutional Neural Network. *Geoderma* **2020**, *367*, 114208. https://doi.org/10.1016/j.geoderma.2020.114208.

42.	Ru, C.; Li, Z.; Tang, R. A Hyperspectral Imaging Approach for Classifying Geographical Origins of Rhizoma Atractylodis Macrocephalae Using the Fusion of Spectrum-Image in VNIR and SWIR Ranges (VNIR-SWIR-FuSI). *Sensors* **2019**, *19*, 2045. https://doi.org/10.3390/s19092045.

43.	Yang, A.; Li, M.; Ding, Y.; Hong, D.; Lv, Y.; He, Y. GTFN: GCN and Transformer Fusion Network With Spatial-Spectral Features for Hyperspectral Image Classification. *IEEE Trans. Geosci. Remote Sens.* **2023**, *61*, 6600115. https://doi.org/10.1109/TGRS.2023.3314616.

44.	Saha, D.; Manickavasagan, A. Machine Learning Techniques for Analysis of Hyperspectral Images to Determine Quality of Food Products: A Review. *Curr. Res. Food Sci.* **2021**, *4*, 28–44.

45.	Zhang, Z.; Li, Y.; Li, C.; Wang, Z.; Chen, Y. Algorithm of Stability-Analysis-Based Feature Selection for NIR Calibration Transfer. *Sensors* **2022**, *22*, 1659. https://doi.org/10.3390/s22041659.

46.	Grimmig, R.; Lindner, S.; Gillemot, P.; Winkler, M.; Witzleben, S. Analyses of Used Engine Oils via Atomic Spectroscopy—Influence of Sample Pre-Treatment and Machine Learning for Engine Type Classification and Lifetime Assessment. *Talanta* **2021**, *232*, 122431. https://doi.org/10.1016/j.talanta.2021.122431.

47.	Hong, D.; Han, Z.; Yao, J.; Gao, L.; Zhang, B.; Plaza, A.; Chanussot, J. SpectralFormer: Rethinking Hyperspectral Image Classification with Transformers. *IEEE Trans. Geosci. Remote Sens.* **2022**, *60*, 5518615. https://doi.org/10.1109/TGRS.2021.3130716.

48.	Peters, R.D.; Noble, S.D. Spectrographic Measurement of Plant Pigments from 300 to 800nm. *Remote Sens. Env.* **2014**, *148*, 119–123. https://doi.org/10.1016/j.rse.2014.03.020.

49.	Bazi, Y.; Bashmal, L.; Al Rahhal, M.M.; Dayil, R.A.; Ajlan, N. Al Vision Transformers for Remote Sensing Image Classification. *Remote Sens.* **2021**, *13*, 516. https://doi.org/10.3390/rs13030516.

50.	Li, H.; Liang, Y.; Xu, Q.; Cao, D. Key Wavelengths Screening Using Competitive Adaptive Reweighted Sampling Method for Multivariate Calibration. *Anal. Chim. Acta* **2009**, *648*, 77–84. https://doi.org/10.1016/j.aca.2009.06.046.

51.	Liu, C.; Chu, Z.; Weng, S.; Zhu, G.; Han, K.; Zhang, Z.; Huang, L.; Zhu, Z.; Zheng, S. Fusion of Electronic Nose and Hyperspectral Imaging for Mutton Freshness Detection Using Input-Modified Convolution Neural Network. *Food Chem.* **2022**, *385*, 132651. https://doi.org/10.1016/j.foodchem.2022.132651.