

## Article

# Application of Attention-Enhanced 1D-CNN Algorithm in Hyperspectral Image and Spectral Fusion Detection of Moisture Content in Orah Mandarin (*Citrus reticulata* Blanco)

Weiqli Li <sup>1</sup> , Yifan Wang <sup>1</sup>, Yue Yu <sup>1</sup> and Jie Liu <sup>1,2,3,\*</sup> 

<sup>1</sup> School of Engineering, Huazhong Agricultural University, Wuhan 430070, China; weiqli\_li@webmail.hzau.edu.cn (W.L.); wyf0706@webmail.hzau.edu.cn (Y.W.); 2023307120007@webmail.hzau.edu.cn (Y.Y.)

<sup>2</sup> Key Laboratory of Agricultural Equipment in Mid-Lower Yangtze River, Ministry of Agriculture and Rural Affairs, Wuhan 430070, China

<sup>3</sup> Citrus Mechanization Research Base, Ministry of Agriculture and Rural Affairs, Wuhan 430070, China

\* Correspondence: liujie@mail.hzau.edu.cn; Tel.: +86-156-2326-3208

**Abstract:** A method fusing spectral and image information with a one-dimensional convolutional neural network (1D-CNN) for the detection of moisture content in Orah mandarin (*Citrus reticulata* Blanco) was proposed. The 1D-CNN model integrated with three different attention modules (SEAM, ECAM, CBAM) and machine learning models were applied to individual spectrum and fused information by passing the traditional feature extraction stage. Additionally, the dimensionality reduction of hyperspectral images and extraction of one-dimensional color and textural features from the reduced images were performed, thus avoiding the large parameter volumes and efficiency decline inherent in the direct modeling of two-dimensional images. The results indicated that the 1D-CNN model with integrated attention modules exhibited clear advantages over machine learning models in handling multi-source information. The optimal machine learning model was determined to be the random forest (RF) model under the fusion information, with a correlation coefficient ( $R$ ) of 0.8770 and a root mean square error (RMSE) of 0.0188 on the prediction set. The CBAM-1D-CNN model under the fusion information exhibited the best performance, with an  $R$  of 0.9172 and an RMSE of 0.0149 on the prediction set. The 1D-CNN models utilizing fusion information exhibited superior performance compared to single spectrum, and 1D-CNN with the fused information based on SEAM, ECAM, and CBAM respectively improved  $R_p$  by 4.54%, 0.18%, and 10.19% compared to the spectrum, with the RMSEP decreased by 11.70%, 14.06%, and 31.02%, respectively. The proposed approach of 1D-CNN integrated attention can obtain excellent regression results by only using one-dimensional data and without feature pre-extracting, reducing the complexity of the models, simplifying the calculation process, and rendering it a promising practical application.

**Keywords:** one-dimensional convolutional neural networks; attention modules; data fusion; moisture content; Orah mandarin



**Citation:** Li, W.; Wang, Y.; Yu, Y.; Liu, J. Application of Attention-Enhanced 1D-CNN Algorithm in Hyperspectral Image and Spectral Fusion Detection of Moisture Content in Orah Mandarin (*Citrus reticulata* Blanco). *Information* **2024**, *15*, 408. <https://doi.org/10.3390/info15070408>

Academic Editor: Marco Leo

Received: 23 May 2024

Revised: 6 July 2024

Accepted: 12 July 2024

Published: 14 July 2024



**Copyright:** © 2024 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

## 1. Introduction

Orah mandarin (*Citrus reticulata* Blanco), a hybrid of Temple orange and Dancy tangerine, combines the flavors of oranges and tangerines and is highly favored by consumers [1]. Its harvesting season spans from December to March of the following year, a typically off-peak period, offering market price advantages and facilitating stable year-round supply. Hence, it is widely cultivated in regions like Guangxi, Yunnan, Sichuan, and Chongqing in China [2]. Moisture content is crucial for Orah mandarins' texture, taste, and freshness. However, Orah mandarins are highly susceptible to natural metabolic processes such as respiration and transpiration after harvest, leading to a decrease in internal organic matter and moisture, thus resulting in a decline in fruit quality and freshness [3,4]. At the same

time, the dry and astringent eating sensation affects consumers' willingness to purchase [5]. Moreover, the lack of quality grading is a significant barrier to high-quality development in China's fruit industry, with only 5% of fresh fruits meeting export standards [6]. With the trend of the thin profit of citrus products, realizing the transformation and upgrading of heavy quality and upgrading with the help of information technology is urgently required [7]. Developing rapid methods for Orah mandarin quality assessment is crucial to meet consumer demands and enhance market competitiveness. With the rise of optical and computational technologies, near-infrared spectroscopy has become prominent in non-destructive detection because of its simple operation, high repeatability, and non-invasive stability, which operates on the principle that molecular bonds selectively absorb light and create unique optical properties that reflect internal chemical and physical information in samples [8,9]. Hyperspectral technology combines digital imaging with spectroscopy to acquire three-dimensional information, including two-dimensional spatial and one-dimensional spectral information [10]. This technology fully utilizes the hyperspectral image information to establish various regression models for the quantitative detection of Orah mandarin [11,12].

To fully exploit the spectral and image information provided by hyperspectral technology and efficiently detect the moisture content of Orah mandarin, this work introduces a method that integrated the image information and the spectrum information in hyperspectral data. Combined with a one-dimensional convolutional neural network (1D-CNN) algorithm, the partial least squares regression (PLSR) algorithm, and machine learning (ML) algorithms, a rapid method to detect the moisture content of the Orah mandarin is determined, providing technical support for the safe development of the Orah mandarin industry.

Here are the highlights of our contributions:

- Performing a linear dimensionality reduction of the hyperspectral image and obtaining its color and textural features.
- Proposing a 1D-CNN model with fusion information of spectral and image information for the detection of moisture content in Orah mandarin and utilizing the powerful information extraction function of the convolutional neural network to reduce information loss.
- Attention mechanisms are integrated into the 1D-CNN model to improve performance, and the addition of the self-attention layer makes the model more focused on the key information of extracted features, which improves accuracy and generalization.
- In particular, visualizing the spatial distribution of moisture content in the sample and dynamically monitoring changing trends at different times, providing a new perspective for the measurement and analysis of the moisture content of Orah mandarin.

The layout of this paper is structured as follows: Section 2 provides a review of existing research on ML and deep learning (DL) techniques for the detection of agricultural products based on hyperspectral images; Section 3 introduces the materials and methods, including the samples and equipment used in the test and a brief description of the data acquisition method and processing method; Section 4 discusses and evaluates the test findings; and Section 5 concludes by summarizing this study's conclusions.

## 2. Related Work

At present, hyperspectral technology combined with chemometrics has been widely used in the field of the quality detection of agricultural products [13]. Jiang et al. distinguished three kinds of *Camellia oleifera* fruits with different mildew degrees and established partial least squares, k-nearest neighbor, and classification regression tree models using a hyperspectral imaging system, in which the accuracy of the least square discriminant analysis (PLS-DA) model in the external prediction set was 90.8% [14]. Xuan et al. [15] obtained hyperspectral images of peach fruit surface, selected effective wavelengths by competitive adaptive reweighted sampling (CARS) and random frog algorithms, and established a multiple linear regression (MLR) model for the rapid evaluation of the

soluble solids content (SSC) and hardness of peach. Cevoli et al. [16] fused visible–near-infrared hyperspectral images and Fourier near-infrared spectroscopy information and established a partial least square model to detect the internal quality parameters of kiwifruit (pulp hardness, soluble solids content, pulp color, and dry matter) to achieve information complementarity among instruments with different spectral ranges. Rajkumar et al. [17] established the relationship between the hyperspectral images and moisture content of banana, and the prediction determination coefficient was 0.87. Gao Sheng et al. [18] established the least square support vector machine (LSSVM) model for the sugar and moisture content of red globe grapes. Hyperspectral images contain spectral and image information, and fully exploiting the advantage of the combination of spectral information will play a greater role in fruit detection. Liu et al. [19] obtained the extracted spectral and image information of hyperspectral images of orange samples for modeling and found that spectral and image fusion information could improve the recognition accuracy of orange in different shelf life, with its misjudgment rate of prediction set to 1.33%. Jiang et al. [20] obtained hyperspectral images of *Agaricus bisporus* with different degrees of damage and extracted spectral features and textural features as inputs. The results showed that the recognition accuracy of the prediction set of the spectral–texture fusion model was 95.56%, better than a single feature input, indicating that multi-source information can improve the detection performance of hyperspectral technology. Convolutional neural networks (CNNs), which use their convolutional layers to extract data-related features and create deeper data representations, play an increasingly greater role in the hyperspectral field [21]. Tian et al. [22] improved the detection performance of the early freezing injury of orange fruit by using a one-dimensional convolution neural network (1D-CNN) model combined with a visible–near-infrared transmission spectrum. Olisah et al. [23] proposed a multi-input convolution neural network model combined with hyperspectral images to detect blackberry maturity feature information, and the prediction accuracy reached 95.1% on external datasets. Chen et al. [24] used an improved CNN to enhance the identification rate of honeysuckle origin based on the near-infrared spectrum. Teng et al. [25] used a 1D-CNN to analyze a complex spectrum.

### 3. Materials and Methods

#### 3.1. Sample Preparation

A total of 195 Orah mandarins from Wuming District in Nanning City of Guangxi with similar size and shape, free from diseases, insect pests, and mechanical damage were selected. Each sample was individually numbered, and then its hyperspectral images were collected under stable temperature conditions.

#### 3.2. Acquisition of Hyperspectral Images

The Zolix Hyper SIS-VNIR-CL hyperspectral imaging system was used to acquire hyperspectral images, which was mainly composed of an electronically controlled mobile platform (Zolix Instruments Co., Ltd., Beijing, China), imaging spectrometer (Spectral Imaging Ltd., Oulu, Finland), halogen tungsten lamp (Zolix Instruments Co., Ltd., Beijing, China), dark box, etc., as shown in Figure 1; its spectral range is 390–1043 nm, and resolution is 2.8 nm, with a total of 520 bands.

#### 3.3. Data Acquisition Method

The process of acquiring the image and spectrum is illustrated in Figure 2. The image acquisition process mainly involved segmenting individual sample hyperspectral images, cropping the region of interest (ROI) of the Orah mandarin to extract the maximum useful information, saving the hyperspectral images of the ROI, and finally extracting color and textural feature parameters from them. The spectrum was obtained through the average spectral reflectance of the ROI.

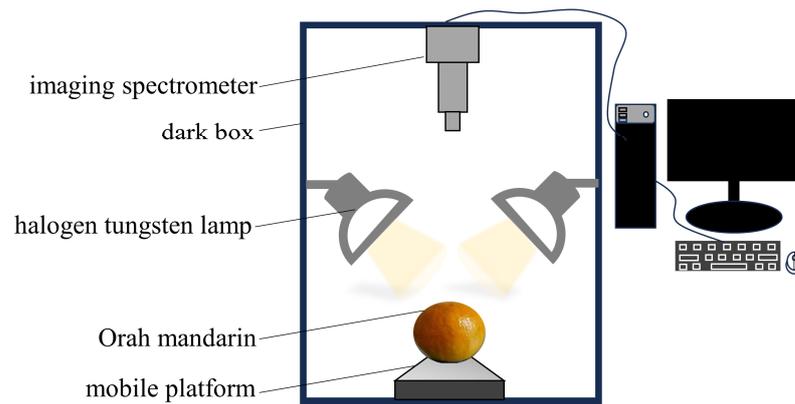


Figure 1. Schematic diagram of hyperspectral imaging system.

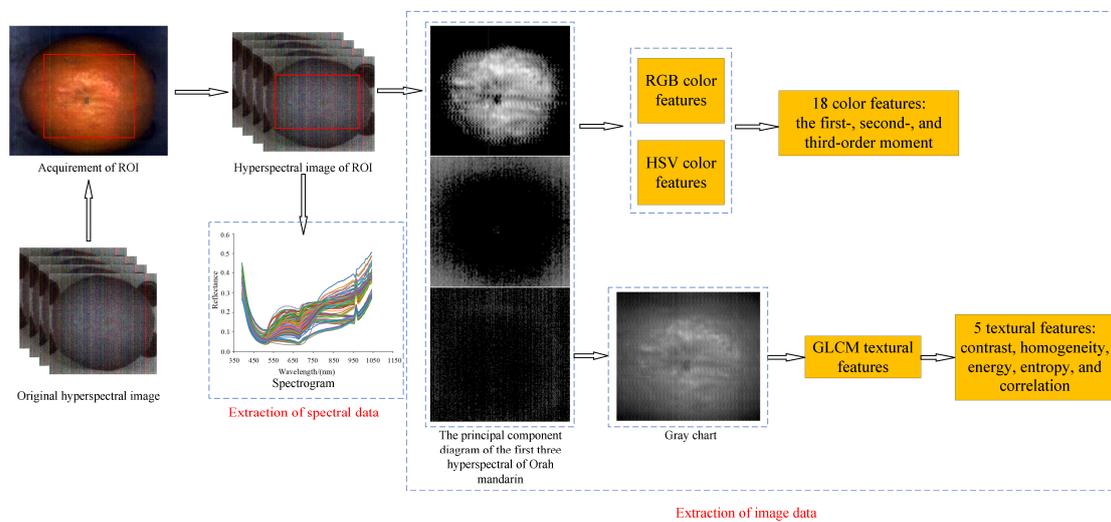


Figure 2. Image and spectral data acquisition process. (In the extraction of spectral data, each color indicated one independent sample).

### 3.3.1. Acquisition and Correction of Hyperspectral Images

The parameters of hyperspectral image acquisition in Orah mandarin were determined by pre-test: exposure time was 20 ms, and acquisition speed was 1.23 mm/s. After preheating for 30 min, each sample fruit pedicel was placed upward in the sample cell to obtain its hyperspectral image. Owing to the influence of dark current and instrument instability, the raw hyperspectral images of the Orah mandarin needed to be collected and corrected by the system’s own software SpectraSENS to eliminate noise. The raw images were corrected using the following Equation (1):

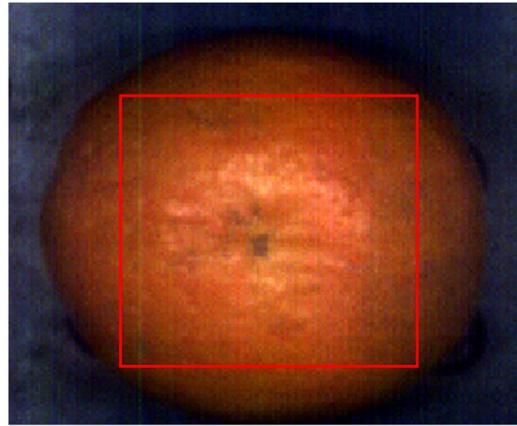
$$I_{ref} = \frac{I - I_d}{I_w - I_d} \times 100\% \tag{1}$$

where  $I$  is the reflection intensity of the raw image,  $I_d$  is the reflection intensity of the collected dark reference image,  $I_w$  is the reflection intensity when the whiteboard is placed, and  $I_{ref}$  is the corrected image reflectivity.

### 3.3.2. The Extraction of the Spectrum, Color, and Texture of Orah mandarin

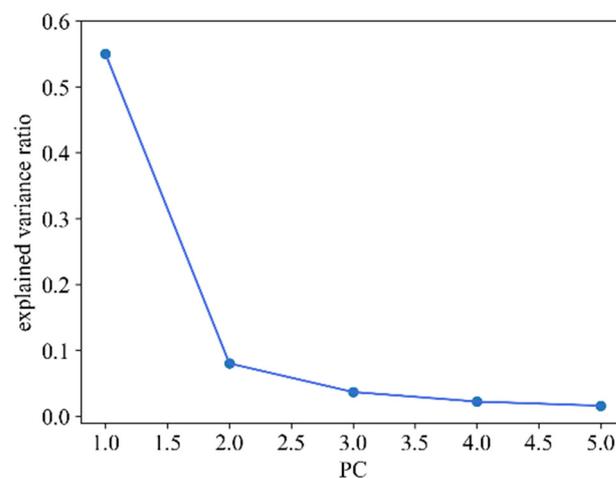
Using the rectangle tool of ENVI5.6 software, the region near the center of each sample was selected as the region of interest (ROI). The selection of the ROI is shown in Figure 3. This method maximizes the cropping of the effective region of Orah mandarin samples, retaining key information while also reducing volume and optimizing the processing workflow. The extraction of the one-dimensional spectrum involved calculating the average

spectral values of all pixels within the ROI of each Orah mandarin, and a total of 195 spectra were extracted. At the same time, the full spectral image of a single sample in the ROI was saved for the subsequent extraction of image texture and color features, and image features consisted of texture and color from principal component images obtained after the dimensionality reduction of ROI hyperspectral images.

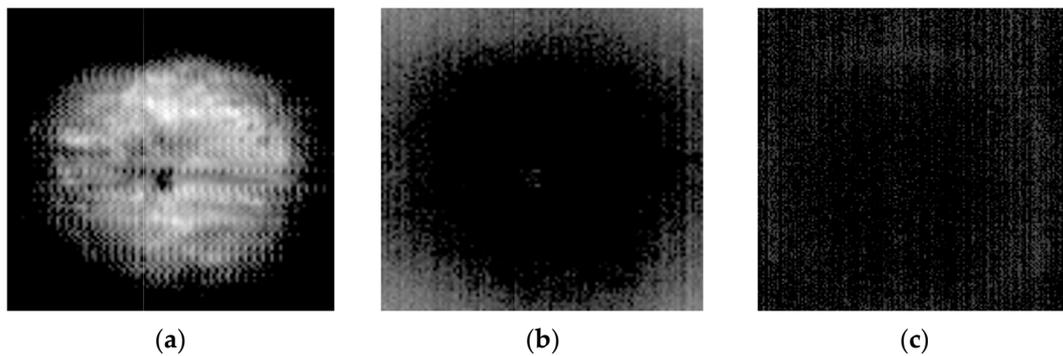


**Figure 3.** Selection of region of interest (ROI) in single sample. (The red square was the ROI).

Principal component analysis (PCA) linearly combines hyperspectral images with different wavelengths into several principal component images, which can effectively solve the problem of high hyperspectral dimensions [26]. Therefore, PCA dimensionality reduction was carried out on the hyperspectral images of 520 channels of Orah mandarin. The variance contribution rate of the principal components is shown in Figure 4. The contribution of the first three principal components was more than 60%, and the subsequent principal components accounted for relatively little. Therefore, the first three principal component images were selected to extract image features, and the first three principal components of Orah mandarin are shown in Figure 5. The RGB and HSV color space contain unique color component information, and the first-, second-, and third-order moments of each channel in the RGB and HSV space were taken as color features [27]. A gray co-occurrence matrix (GLCM) was used to extract the textural features of the gray image converted from the principal component images, and five independent textural feature parameters including contrast, homogeneity, energy, entropy, and correlation were used [28,29]. In summary, a total of 23 image features were extracted, including 18 color features and 5 textural features. The dimensional effect of the different types of datasets was eliminated after normalization.



**Figure 4.** The explainable variance contribution ratio of the first five principal components.



**Figure 5.** First three principal components: (a) the first principal component; (b) the second principal component; and (c) the third principal component.

### 3.3.3. Data Fusion of Spectral and Image Data

The data fusion discussed in this work involved integrating a single spectrum with single image information at the data layer, where different types of datasets were directly concatenated. Concatenated datasets contained much richer information than the single spectral dataset. For a single Orah mandarin, merging spectral and image information generated 543 variables, with contributions of 520 variables from spectrum and 23 variables from image features.

### 3.4. Reference Value Measurement of Moisture Content

According to the GB5009.3-2016 [30] method, 3 pieces of non-adjacent pulp of Orah mandarin were taken as the measuring object, and the mass  $m_1$  was obtained by weighing in an aluminum box (mass  $m_0$ ), and then they were dried in 101-00B electrothermal constant temperature blast drying oven (Mingtuo Machinery equipment Co., Ltd., Xuchang, China) at a set temperature of 105 °C. The mass of the dried sample and aluminum box was  $m_2$ , and the moisture content was calculated according to the following Equation (2)

$$W = \frac{m_1 - m_2}{m_1 - m_0} \times 100\% \quad (2)$$

where  $W$  is the moisture content of the sample (%),  $m_0$  is the mass of the aluminum box (mg),  $m_1$  is the mass of the aluminum box and pulp before drying (mg), and  $m_2$  is the mass of the aluminum box and pulp after drying (mg).

### 3.5. Data Preprocessing

In this study, SG smoothing was used to improve the signal-to-noise ratio of the spectrum, and the curve was smoothed on the basis of retaining the original spectrum. At the same time, the maximum and minimum normalization method was used to reduce the size gap caused by different dimensions of different variables [31]. The maximum and minimum normalization were calculated using the following Equation (3):

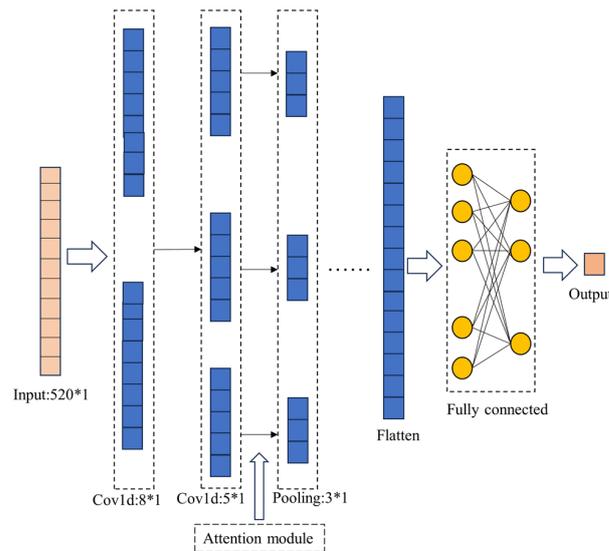
$$X'_{m,n} = \frac{X_{m,n} - \min(X_n)}{\max(X_n) - \min(X_n)} \quad (3)$$

where  $X'_{m,n}$  represents the normalized data of  $m$  row and  $n$  column,  $X_{m,n}$  represents the data of  $m$  row and  $n$  column, and  $X_n$  represents the data of the  $n$  column of the original data.

### 3.6. Model Establishment Methods of 1D-CNN

Convolution neural networks (CNNs) have the ability to deal with complex information automatically. Recent studies have found that the shallow CNN model (with 1–2 convolution layers) has better performance than traditional stoichiometric methods [32]. In this study, a 1D-CNN model based on a hyperspectral image was proposed, which consisted of one input layer, two convolution layers, one pooling layer, one fully connected

layer, and one output layer. By reducing the number of network parameters to enhance the calculation rate and prevent overfitting, its structure is shown in Figure 6. In the 1D-CNN structure, the convolution layer extracted the features related to the moisture content of Orah mandarin, and three different attention modules, including the squeeze-and-excitation attention module (SEAM), efficient channel attention module (ECAM), and convolutional block attention module (CBAM), were added after the convolution layer to enhance the ability of feature representation, and the maximum pooling layer was used to reduce the computational complexity and enhance feature invariance. Finally, the fully connected layer neurons analyzed and processed the inputs globally and mapped them to the output space.



**Figure 6.** Structure of attention-enhanced 1D-CNN.

The SEAM compresses spatial features of each channel into a scalar through average pooling and generates channel weights through a fully connected layer, which enhances important features while attenuating unimportant ones [33]. The ECAM utilizes convolution layers to generate channel weights, avoiding the computational overhead associated with fully connected layers [34]. The CBAM combines channel attention and spatial attention to enhance feature representation. Channel features are extracted using average pooling and maximum pooling, and attention weights are generated through fully connected networks, which are then applied to recalibrate the original feature map of channels. The spatial attention mechanism conducts average pooling and maximum pooling operations on the calibrated feature map to obtain two-dimensional feature maps. And then, the spatial attention weights obtained through convolution operations are then applied to the feature map of spatial positions calibrated by the channel attention mechanism [35].

The convolutional layer of the 1D-CNN was used to extract features from the feature vector output of the previous layer. By stacking multiple convolutional layers, deeper abstract features that were more conducive to regression can be learned. The output of the convolutional layer was calculated using the following Equation (4):

$$x_j^l = f\left(\sum_{i=1}^m x_i^{l-1} * k_{ij}^l + b_j^l\right) \quad (4)$$

where  $x_j^l$  is the  $j$ <sub>th</sub> feature vector output of the  $l$ <sub>th</sub> layer,  $k_{ij}^l$  and  $b_j^l$  are the convolutional kernel weights and bias of the  $l$ <sub>th</sub> layer, respectively,  $m$  is the number of output feature vectors in the  $l$ <sub>th</sub> layer,  $*$  denotes the convolution operator, and  $f$  is the activation function. The activation function was set as the rectified linear unit (RELU) in this study, and the function expression of the RELU is shown in Equation (5).

$$f(x) = \max(0, x) \quad (5)$$

where  $x$  is the feature vector,  $\max$  denotes the function that computes the maximum value, and  $f$  is the RELU function.

### 3.7. Model Evaluation Metrics

The correlation coefficient ( $R$ ) and root mean square error (RMSE) were used as the performance evaluation metrics of the model, in which  $R$  reflected the fitting degree of predicted values and reference values, and RMSE represented the deviation between the predicted value and the reference value. The above parameters were calculated by Equations (6) and (7):

$$R = \frac{\sum_{i=1}^n [(\hat{y}_i - \bar{y}_i)(y_i - \bar{y}_i)]}{\sqrt{\sum (\hat{y}_i - \bar{y}_i)^2 \times \sum (y_i - \bar{y}_i)^2}} \quad (6)$$

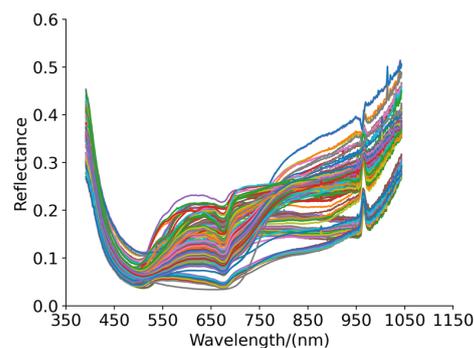
$$RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^n (\hat{y}_i - y_i)^2} \quad (7)$$

where  $n$  is the number of samples in the dataset,  $\hat{y}_i$  is the predicted value of the moisture content of sample  $I$ ,  $y_i$  is the measured value of the moisture content of sample  $I$ , and  $\bar{y}_i$  is the average moisture content of  $n$  samples.

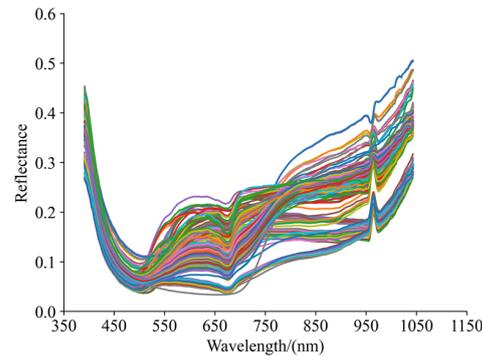
## 4. Results and Discussion

### 4.1. Spectral Analysis and Preprocessing of Orah Mandarin

The average spectral curves of 195 Orah mandarin samples are presented in Figure 6. All sample curves exhibit consistent trends, indicating that the absorption characteristics of the spectra are primarily determined by common internal tissue properties of Orah mandarin. However, differences in reflectance intensity reflect variations in individual internal chemical composition. It is clear that in Figure 7, there is no significant difference in reflectivity intensity among different samples in the range of 390–550 nm, but there are great differences in spectral intensity among different samples in the range of 550–1043 nm with crosses and overlaps between spectra, which is necessary to combine spectral and stoichiometric analysis. The valley observed near 500 nm is mainly related to carotenoids and other substances in Orah mandarin [36]. The strong absorption peak near 670 nm is the effective wavelength of chlorophyll, while the peak near 530 nm is related to the absorption bands of chlorophyll and carotenoids on the Orah mandarin surface [19]. At approximately 960 nm, there is a tensile vibration of water or carbohydrates in the third-order frequency doubling of the O-H bond [37]. Because the smoothness of the curves is within the range of 540–760 nm, while noise is more pronounced at both ends, the original spectra were subjected to Savitzky–Golay (S-G) smoothing, in which the number of polynomials was set at 3, with a window size of 5 in the first segment (390–540 nm) and a window size of 12 in the tail segment (760–1043 nm). The filtered spectrum is shown in Figure 8, which shows a spectral curve that reduced noise and retained the original information.



**Figure 7.** Origin spectra of Orah mandarin. (Each color indicated one independent sample).



**Figure 8.** Spectra of Orah mandarin after SG. (Each color indicated one independent sample).

#### 4.2. Reference Measurements

The range, mean, and standard deviation values of the moisture content of the Orah mandarin were analyzed and are presented in Table 1. The moisture content of 195 samples varied from 6.10% to 8.95%, with a mean of 8.03% and a standard deviation of 0.61%.

**Table 1.** Descriptive statistics for moisture content of all samples.

| Quality Index        | Sample Set     | Samples | Maximum Value | Minimum Value | Average Value | Standard Deviation |
|----------------------|----------------|---------|---------------|---------------|---------------|--------------------|
| Moisture content (%) | All samples    | 195     | 91.82%        | 75.07%        | 84.06%        | 3.67%              |
|                      | Correction set | 146     | 91.82%        | 75.07%        | 84.07%        | 3.61%              |
|                      | Prediction set | 49      | 89.27%        | 77.08%        | 84.01%        | 3.40%              |

For reliability and stability of the model, all samples were divided into a correction set of 146 samples and a prediction set of 49 samples using the KS (Kennard–Stone) algorithm [38], and the statistics for moisture content are shown in the last two rows of Table 1. The coverage of the correction set was larger than that of the prediction set during the modeling process and may contribute to a robust calibration model.

#### 4.3. Establishment and Analysis of Quantitative Model

To compare the modeling effects of traditional machine learning algorithms and the 1D-CNN algorithm under the spectral and fusion information of the concatenation of spectral and image features, partial least squares regression (PLSR), support vector regression (SVR), artificial neural network (ANN), decision Tree (DT), random forest (RF), light gradient boosting machine (LightGBM), and 1D-CNN models were established, and the performances of the models were compared by the correlation coefficient of the prediction set ( $R_p$ ) and prediction root mean square error of the prediction set (RMSEP). Additionally, it is desirable that the correlation coefficient of the correction set ( $R_c$ ) and root mean square error of the correction set (RMSEC) do not differ significantly from those of the prediction set, as this indicates good model stability and suggests that the model can effectively predict unknown samples. A total of 520 spectral bands and 23 image features were normalized to 0 and 1 before modeling. The concatenated fused information, comprising a total of 543 feature variables, was then used as the modeling input.

##### 4.3.1. Establishment and Analysis of PLSR and Machine Learning Model

The results obtained for predicting the dataset based on different detection models are shown in Table 2. The modeling results indicated that the incorporation of image information improves the moisture content detection performance of the PLSR, SVR, ANN, DT, and RF models. Among them, the RF model under spectral fusion achieved the best performance, with an  $R_p$  of 0.8770 and an RMSEP of 0.0188 on the prediction set. However, the lightGBM model was insensitive to the input of image textural and color features, resulting in lower accuracy compared to the single spectrum.

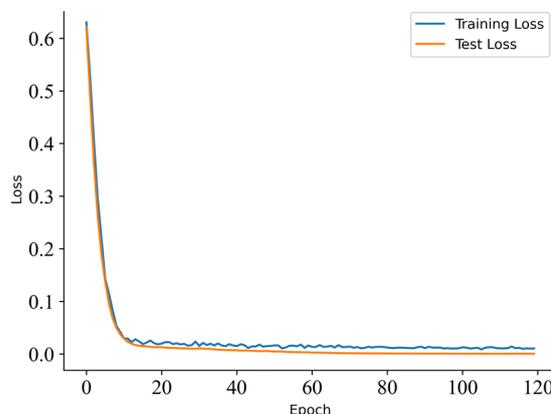
**Table 2.** The parameters of the moisture content prediction models of spectral and fusion information based on different machine learning algorithms.

| Modeling Data | Input Number | Model    | Correction Set |        | Prediction Set |        |
|---------------|--------------|----------|----------------|--------|----------------|--------|
|               |              |          | $R_C$          | RMSEC  | $R_p$          | RMSEP  |
| Spectral data | 520          | PLSR     | 0.8572         | 0.0185 | 0.8535         | 0.0180 |
|               |              | SVR      | 0.8747         | 0.0174 | 0.8547         | 0.0179 |
|               |              | ANN      | 0.6757         | 0.0248 | 0.5696         | 0.0205 |
|               |              | DT       | 0.8757         | 0.0173 | 0.8410         | 0.0187 |
|               |              | RF       | 0.9446         | 0.0122 | 0.8227         | 0.0196 |
|               |              | LightGBM | 0.8992         | 0.0162 | 0.8535         | 0.0182 |
| Fused data    | 543          | PLSR     | 0.8781         | 0.0169 | 0.8718         | 0.0178 |
|               |              | SVR      | 0.8681         | 0.0175 | 0.8752         | 0.0198 |
|               |              | ANN      | 0.7712         | 0.0227 | 0.7557         | 0.0246 |
|               |              | DT       | 0.8713         | 0.0173 | 0.8629         | 0.0184 |
|               |              | RF       | 0.9267         | 0.0137 | 0.8770         | 0.0188 |
|               |              | LightGBM | 0.8987         | 0.0174 | 0.7524         | 0.0252 |

#### 4.3.2. Establishment and Analysis of Attention-Enhanced 1D-CNN Model

The CNN model of moisture content based on single spectral input and fusion information input was composed of two convolution layers and one max pooling layer with a size of  $3 \times 1$ . The number of convolution nuclei in the 1D-CNN model of a single spectrum was set at 8 and 16, and the size was  $8 \times 1$  and  $3 \times 1$ , respectively. The number of convolution kernels in the fused feature CNN model was set at 8 and 16, respectively, and the size of kernels was  $3 \times 1$ . The activation function was set as the rectified linear unit (ReLU), and the loss function was defined as the mean square error (MSE) loss.

Table 3 presents the regression results obtained by a 1D-CNN model integrating different attention mechanism modules using the single spectral input and fusion information input of hyperspectral images of Orah mandarin. Overall, the 1D-CNN model using fusion information input consistently performed better than the single spectrum, among which the fused features based on the SEAM, ECAM, and CBAM improved the  $R_p$  by 4.54%, 0.18%, and 10.19%, respectively, compared with the single-spectrum modeling, decreasing the RMSEP by 11.70%, 14.06%, and 31.02%, respectively. CBAM-1D-CNN with fusion information input had the best results, with an  $R_p$  of 0.9172 and RMSEP of 0.0149, and its loss function curve is shown in Figure 9 with the loss value tending to stabilize after a certain number of epochs, and its prediction set results are shown in Figure 10 with a coincidence of the prediction values and measured value. This highlights the advantages of the 1D-CNN model in learning and combining relevant features from different sources and improving its detection capabilities.



**Figure 9.** Loss curves of CBAM-1D-CNN model based on fused input.

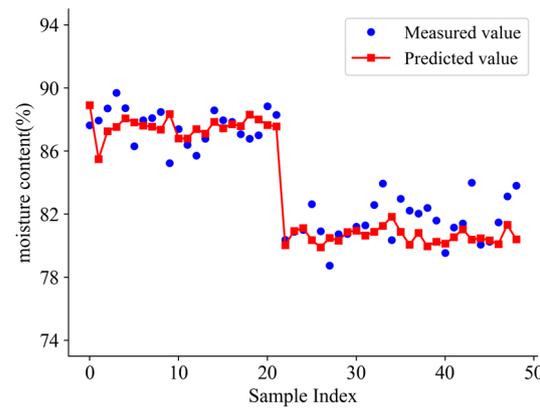


Figure 10. Results of CBAM-1D-CNN based on fused input on prediction set.

Table 3. The parameters of the moisture content prediction models of spectral and fusion information based on the attention-enhanced 1D-CNN.

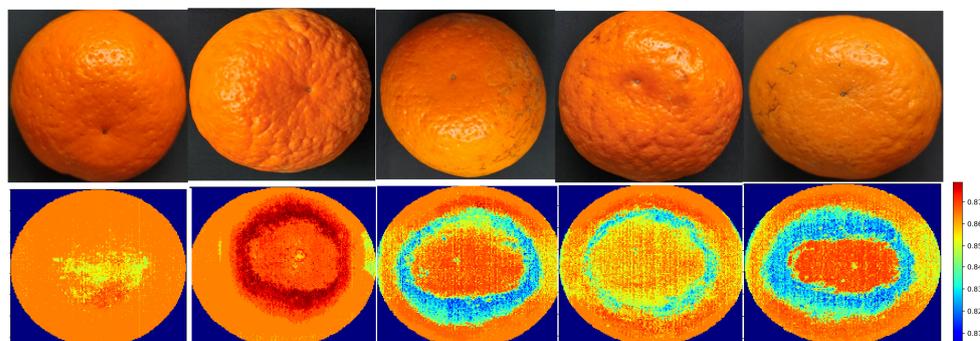
| Modeling Data | Input Number | Model         | Correction Set |        | Prediction Set |        |
|---------------|--------------|---------------|----------------|--------|----------------|--------|
|               |              |               | $R_C$          | RMSEC  | $R_p$          | RMSEP  |
| Spectral data | 520          | SEAM + 1D-CNN | 0.8573         | 0.0201 | 0.8614         | 0.0170 |
|               |              | ECAM + 1D-CNN | 0.8545         | 0.0232 | 0.8831         | 0.0192 |
|               |              | CBAM + 1D-CNN | 0.8255         | 0.0241 | 0.8324         | 0.0216 |
| Fused data    | 543          | SEAM + 1D-CNN | 0.8403         | 0.0202 | 0.9005         | 0.0150 |
|               |              | ECAM + 1D-CNN | 0.8369         | 0.0212 | 0.8847         | 0.0165 |
|               |              | CBAM + 1D-CNN | 0.8609         | 0.0192 | 0.9172         | 0.0149 |

The above experimental results indicated that the 1D-CNN model integrated with the attention mechanism module can detect the moisture content of Orah mandarin by using one-dimensional data, and its performance in multi-source data fusion was better than the traditional machine learning methods. The detection performance of the 1D-CNN model integrated with the CBAM was further enhanced. The CBAM utilizes a dual attention mechanism to comprehensively improve the feature representation capability, enabling the model to better capture important spectral and image data. The CBAM-1DCNN model demonstrated improvements in the correlation coefficient and root mean square error of prediction over the PLSR, SVR, ANN, DT, RF, and LightGBM models with fusion information, achieving increases by 5.21%, 4.80%, 21.37%, 6.29%, 4.58%, and 21.90% in  $R_p$ , respectively, and reductions by 16.29%, 24.75%, 39.43%, 19.02%, 20.74%, and 40.87% in the RMSEP. In this study, only one-dimensional features were used in the algorithm construction process without the need to extract features in advance, which helps to reduce the number of parameters and simplify the calculation steps.

#### 4.4. Visual Distribution of Moisture Content in Orah Mandarin

To visualize and analyze the moisture content of Orah mandarin, an in-depth exploration of their internal distribution characteristics was conducted. The raw hyperspectral images of the same batch of samples were collected every three days, resulting in a total of five collections of hyperspectral images in a continuous storage period. An optimal 1D-CNN model based on spectrum was utilized to predict the moisture content values of Orah mandarin at individual pixel points. Combining color mapping and dynamic color adjustment based on minimum and maximum values, visualization distribution maps of the moisture content of Orah mandarin were obtained. The true-color map and the pseudo-color distribution of moisture content of Orah mandarin are shown in Figure 11, representing images of Orah mandarin with increasing storage periods from left to right. The top row shows true-color maps, while the bottom row displays pseudo-color maps. In the pseudo-color maps, the transition from blue to red indicated an increase in moisture

content, effectively demonstrating the spatial distribution of moisture content in Orah mandarin. Overall, the Orah mandarin exhibited relatively high moisture content for appearing in shades of orange-red. Furthermore, the moisture content of Orah mandarin showed a pattern from a relatively uniform overall distribution to gradually expanding regions of decreasing moisture content in the end of the storage period, providing important references for further data interpretation and research.



**Figure 11.** A true-color and visual distribution map of the moisture content of Orah mandarin during a period.

## 5. Conclusions

Hyperspectral images (390–1043 nm) and the 1D-CNN algorithm were used to detect the moisture content of Orah mandarin, and PLSR, multiple machine learning, and 1D-CNN models were applied to spectral and fused features. The results showed that the 1D-CNN model with the integrated attention mechanism module was superior to PLSR and machine learning models in processing multi-source information. Compared with PLSR, SVR, ANN, DT, RF, and LightGBM with fusion information inputs, CBAM-1DCNN increased the correlation coefficient  $R_p$  by 5.21%, 4.80%, 21.37%, 6.29%, 4.58%, and 21.90%, respectively, and decreased the RMSEP by 16.29%, 24.75%, 39.43%, 19.02%, 20.74%, and 40.87%, demonstrating the advantages of 1D-CNN in processing the complexity and interaction of color and textural features in hyperspectral images. This method of the 1D-CNN utilized one-dimensional features exclusively and did not require feature extraction beforehand to effectively reduce parameter volume and simplify computational steps, which was considered as a promising practice application choice. By combining the pseudo-color distribution of moisture content with hyperspectral images of Orah mandarin, intuitive information regarding the moisture content distribution was provided. The distribution graph of moisture content over a storage time clearly illustrated a trend from generally high moisture content to a continuous decrease, which held significant guiding implications for the high-quality development of the citrus industry.

**Author Contributions:** Conceptualization, W.L. and Y.W.; methodology, J.L. and W.L.; software, W.L.; validation, W.L., Y.W., and Y.Y.; formal analysis, W.L.; investigation, W.L.; resources, J.L.; data curation, Y.W.; writing—original draft preparation, W.L.; writing—review and editing, J.L. and W.L.; visualization, Y.Y. and W.L. All authors have read and agreed to the published version of the manuscript.

**Funding:** This research was funded by National Natural Science Foundation of China, grant number 61205153.

**Institutional Review Board Statement:** Not applicable.

**Informed Consent Statement:** Not applicable.

**Data Availability Statement:** The data presented in this study are available on request from the corresponding author due to privacy or ethical restrictions.

**Conflicts of Interest:** The authors declare no conflicts of interest.

## Abbreviations

The following abbreviations are used in this manuscript:

|          |  |
|----------|--|
| ML       | machine learning                             |
| DL       | deep learning                                |
| 1D-CNN   | one-dimensional convolutional neural network |
| ECAM     | squeeze-and-excitation attention module      |
| ECAM     | efficient channel attention module           |
| CBAM     | convolutional block attention module         |
| PLSR     | partial least squares regression             |
| SVR      | support vector regression                    |
| ANN      | artificial neural network                    |
| DT       | decision Tree                                |
| RF       | random forest                                |
| LightGBM | light gradient boosting machine              |
| ROI      | region of interest                           |

## References

- Huang, Q.; Liu, J.; Hu, C.; Wang, N.; Zhang, L.; Mo, X.; Li, G.; Liao, H.; Huang, H.; Ji, S. Integrative analyses of transcriptome and carotenoids profiling revealed molecular insight into variations in fruits color of *Citrus reticulata* Blanco induced by transplantation. *Genomics* **2022**, *114*, 110291. [[CrossRef](#)] [[PubMed](#)]
- He, Y.; Li, W.; Zhu, P.; Wang, M.; Qiu, J.; Sun, H.; Zhang, R.; Liu, P.; Ling, L.; Fu, X. Comparison between the vegetative and fruit characteristics of ‘Orah’ (*Citrus reticulata* Blanco) mandarin under different climatic conditions. *Sci. Hortic.* **2022**, *300*, 111064. [[CrossRef](#)]
- Maurizzi, E.; Bigi, F.; Volpelli, L.A.; Pulvirenti, A. Improving the post-harvest quality of fruits during storage through edible packaging based on guar gum and hydroxypropyl methylcellulose. *Food Packag. Shelf Life* **2023**, *40*, 101178. [[CrossRef](#)]
- Dhami, K.; Asrey, R.; Awasthi, O.; Bhowmik, A. Pre and postharvest treatments of methyl jasmonate: Maintain quality and shelf-life of Kinnow mandarin fruit during cold storage. *South Afr. J. Bot.* **2022**, *151*, 808–815. [[CrossRef](#)]
- Yuan, Y.; Chen, X. Vegetable and fruit freshness detection based on deep features and principal component analysis. *Curr. Res. Food Sci.* **2024**, *8*, 100656. [[CrossRef](#)]
- Wang, H.; Zhou, Z.; Yang, Z.; Cao, Y.; Zhang, C. Constraints on the high-quality development of Chinese fruit industry. *China Fruits* **2023**, *65*, 1–9. [[CrossRef](#)]
- Chen, H.; Wu, Q.; Xu, B.; Wang, B.; Jiang, H. Exploration and practice of citrus information system. *China Fruits* **2024**, *66*, 122–127. [[CrossRef](#)]
- Mei, M.; Li, J. An overview on optical non-destructive detection of bruises in fruit: Technology, method, application, challenge and trend. *Comput. Electron. Agric.* **2023**, *213*, 108195. [[CrossRef](#)]
- Wang, M.; Xu, Y.; Yang, Y.; Mu, B.; Nikitina, M.A.; Xiao, X. Vis/NIR optical biosensors applications for fruit monitoring. *Biosens. Bioelectron. X* **2022**, *11*, 100197. [[CrossRef](#)]
- Min, D.; Zhao, J.; Bodner, G.; Ali, M.; Li, F.; Zhang, X.; Rewald, B. Early decay detection in fruit by hyperspectral imaging—Principles and application potential. *Food Control* **2023**, *152*, 109830. [[CrossRef](#)]
- Wakchaure, G.; Nikam, S.B.; Barge, K.R.; Kumar, S.; Meena, K.K.; Nagalkar, V.J.; Choudhari, J.; Kad, V.; Reddy, K.S. Maturity stages detection prototype device for classifying custard apple (*Annona squamosa* L) fruit using image processing approach. *Smart Agric. Technol.* **2024**, *7*, 100394. [[CrossRef](#)]
- Rodríguez-Ortega, A.; Aleixos, N.; Blasco, J.; Albert, F.; Munera, S. Study of light penetration depth of a Vis-NIR hyperspectral imaging system for the assessment of fruit quality. A case study in persimmon fruit. *J. Food Eng.* **2023**, *358*, 111673. [[CrossRef](#)]
- Sanchez, P.D.C.; Hashim, N.; Shamsudin, R.; Nor, M.Z.M. Applications of imaging and spectroscopy techniques for non-destructive quality evaluation of potatoes and sweet potatoes: A review. *Trends Food Sci. Technol.* **2020**, *96*, 208–221. [[CrossRef](#)]
- Jiang, H.; Jiang, X.; Ru, Y.; Chen, Q.; Li, X.; Xu, L.; Zhou, H.; Shi, M. Rapid and non-destructive detection of natural mildew degree of postharvest *Camellia oleifera* fruit based on hyperspectral imaging. *Infrared Phys. Technol.* **2022**, *123*, 104169. [[CrossRef](#)]
- Xuan, G.; Gao, C.; Shao, Y. Spectral and image analysis of hyperspectral data for internal and external quality assessment of peach fruit. *Spectrochim. Acta Part A Mol. Biomol. Spectrosc.* **2022**, *272*, 121016. [[CrossRef](#)] [[PubMed](#)]
- Cevoli, C.; Iaccheri, E.; Fabbri, A.; Ragni, L. Data fusion of FT-NIR spectroscopy and Vis/NIR hyperspectral imaging to predict quality parameters of yellow flesh “Jintao” kiwifruit. *Biosyst. Eng.* **2024**, *237*, 157–169. [[CrossRef](#)]
- Rajkumar, P.; Wang, N.; Elmasry, G.; Raghavan, G.; Garipey, Y. Studies on banana fruit quality and maturity stages using hyperspectral imaging. *J. Food Eng.* **2012**, *108*, 194–200. [[CrossRef](#)]
- Gao, S.; Wang, Q. Non-destructive testing of red globe grape sugar content and moisture content based on visible/near infrared spectroscopy transmission technology. *Chin. Opt.* **2021**, *14*, 566–577.
- Liu, Y.; Wang, S. Research on Non-Destructive Testing of Navel Orange Shelf-Life Imaging Based on Hyperspectral Image and Spectrum Fusion. *Spectrosc. Spectr. Anal.* **2022**, *42*, 1792–1797.

20. Jiang, F.; Shen, Z.; Yang, L.; Chen, Y.; Sun, B. Hyperspectral imaging for quality detection of *Agaricus bisporus* under vibration stress. *Food Sci.* **2022**, *43*, 353–360. [[CrossRef](#)]
21. Wieme, J.; Mollazade, K.; Malounas, I.; Zude-Sasse, M.; Zhao, M.; Gowen, A.; Argyropoulos, D.; Fountas, S.; Van Beek, J. Application of hyperspectral imaging systems and artificial intelligence for quality assessment of fruit, vegetables and mushrooms: A review. *Biosyst. Eng.* **2022**, *222*, 156–176. [[CrossRef](#)]
22. Tian, S.; Wang, S.; Xu, H. Early detection of freezing damage in oranges by online Vis/NIR transmission coupled with diameter correction method and deep 1D-CNN. *Comput. Electron. Agric.* **2022**, *193*, 106638. [[CrossRef](#)]
23. Olisah, C.C.; Trehwella, B.; Li, B.; Smith, M.L.; Winstone, B.; Whitfield, E.C.; Fernández, F.F.; Duncalfe, H. Convolutional neural network ensemble learning for hyperspectral imaging-based blackberry fruit ripeness detection in uncontrolled farm environment. *Eng. Appl. Artif. Intell.* **2024**, *132*, 107945. [[CrossRef](#)]
24. Chen, D.; Zhang, H.; Lin, L.; Zhang, Z.; Zeng, J.; Chen, L.; Chen, X. Auto-encoder design based on the 1D-VD-CNN model for the detection of honeysuckle from unknown origin. *J. Pharm. Biomed. Anal.* **2023**, *234*, 115572. [[CrossRef](#)] [[PubMed](#)]
25. Teng, Y.; Chen, Y.; Chen, X.; Zuo, S.; Li, X.; Pan, Z.; Shao, K.; Du, J.; Li, Z. Revealing the adulteration of sesame oil products by portable Raman spectrometer and 1D CNN vector regression: A comparative study with chemometrics and colorimetry. *Food Chem.* **2024**, *436*, 137694. [[CrossRef](#)]
26. Li, J.; Luo, W.; Wang, Z.; Fan, S. Early detection of decay on apples using hyperspectral reflectance imaging combining both principal component analysis and improved watershed segmentation method. *Postharvest Biol. Technol.* **2019**, *149*, 235–246. [[CrossRef](#)]
27. Jiang, X.; Tian, J.; Huang, H.; Hu, X.; Han, L.; Huang, D.; Luo, H. Nondestructive visualization and quantification of total acid and reducing sugar contents in fermented grains by combining spectral and color data through hyperspectral imaging. *Food Chem.* **2022**, *386*, 132779. [[CrossRef](#)]
28. Wang, Z.; Fan, S.; An, T.; Zhang, C.; Chen, L.; Huang, W. Detection of Insect-Damaged maize seeds using hyperspectral imaging and hybrid 1D-CNN-BiLSTM models. *Infrared Phys. Technol.* **2024**, *137*, 105208. [[CrossRef](#)]
29. Qiao, M.; Cui, T.; Xia, G.; Xu, Y.; Li, Y.; Fan, C.; Han, S.; Dong, J. Integration of spectral and image features of hyperspectral imaging for quantitative determination of protein and starch contents in maize kernels. *Comput. Electron. Agric.* **2024**, *218*, 108718. [[CrossRef](#)]
30. GB/T 5009.3-2016; National Standard for Food Safety Determination of Moisture in Foods. National Health and Family Planning Commission of the People's Republic of China: Beijing, China, 2016.
31. Hu, R.; Yu, Y.; Ni, M.; Yu, J.; Zhou, J. Identification of Lotus Seed Flour Adulteration Based on Near-Infrared Spectroscopy Combined with Deep Belief Network. *Food Sci.* **2020**, *41*, 298–303. [[CrossRef](#)]
32. Luo, N.; Xu, D.; Xing, B.; Yang, X.; Sun, C. Principles and applications of convolutional neural network for spectral analysis in food quality evaluation: A review. *J. Food Compos. Anal.* **2024**, *128*, 105996. [[CrossRef](#)]
33. Hu, J.; Shen, L.; Sun, G. Squeeze-and-excitation networks. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Salt Lake City, UT, USA, 18–23 June 2018; pp. 7132–7141.
34. Wang, Q.; Wu, B.; Zhu, P.; Li, P.; Zuo, W.; Hu, Q. ECA-Net: Efficient channel attention for deep convolutional neural networks. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, Seattle, WA, USA, 14–19 June 2020; pp. 11534–11542.
35. Woo, S.; Park, J.; Lee, J.-Y.; Kweon, I.S. CBAM: Convolutional block attention module. In Proceedings of the European Conference on Computer Vision (ECCV), Munich, Germany, 8–14 September 2018; pp. 3–19.
36. Liu, Y.; Li, M.; Hu, J.; Xu, Z.; Cui, H. Detection of Citrus Granulation Based on Near-Infrared Hyperspectral Data. *Spectrosc. Spectr. Anal.* **2022**, *42*, 1366–1371. [[CrossRef](#)]
37. Zhang, Y.; Lee, W.S.; Li, M.; Zheng, L.; Ritenour, M.A. Non-destructive recognition and classification of citrus fruit blemishes based on ant colony optimized spectral information. *Postharvest Biol. Technol.* **2018**, *143*, 119–128. [[CrossRef](#)]
38. Andries, J.P.; Vander Heyden, Y. Calibration set reduction by the selection of a subset containing the best fitting samples showing optimally predictive ability. *Talanta* **2024**, *266*, 124943. [[CrossRef](#)]

**Disclaimer/Publisher's Note:** The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.