



OPEN VGG-MFO-orange for sweetness prediction of Linhai mandarin oranges

Chun Fang¹, Runhong Shen², Meiling Yuan¹, ZhengXu¹, Wangyi Ye¹, Sheng Dai¹ & Di Wang³✉

Mandarin orange is a popular fruit in China and known worldwide for its unique flavor and nutritional benefits. As consumer demand for fruit quality increases, the fine assessment and grading of fruit sweetness—especially through non-destructive testing techniques—are becoming increasingly important in agriculture and commerce. In this paper, a new Attention for Orange (AO) attention mechanism and Multiscale Feature Optimization (MFO) feature extraction module are designed and combined with VGG13 convolutional neural network (CNN), innovatively proposed VGG-MFO-Orange CNN model for accurately classifying mandarin oranges with different sweetness. First, a sample of Linhai mandarin oranges was collected, and a sweetness triple classification dataset with 5022 images was formed, utilizing image acquisition and sugar detection. The proposed model was then trained against six influential classical CNN models: DenseNet121, MobileNet_v2, ResNet50, ShuffleNet, VGG13, and VGG13_bn. The experimental results showed that our model achieved an accuracy of 86.8% on the validation set, which was significantly better than the other six models. It also demonstrated excellent generalization ability and effectiveness in predicting the sweetness of Linhai mandarin oranges. Therefore, our model can provide an efficient means of fruit grading for agricultural production, contribute to agricultural modernization, and enhance the competitiveness of agricultural products in the market.

Keywords Linhai mandarin orange, Sugar testing, Deep learning, Convolutional neural networks, Quality classification

Linhai mandarin orange, cultivated along the southeast coast of Zhejiang Province, is an important agricultural product with an annual output of up to 300,000 tons. It has been granted China's National Geographical Indication for Agricultural Products due to its distinctive quality¹. The development of rapid and efficient quality testing techniques is essential for controlling the quality of mandarin oranges, meeting consumer demand for quality fruits, and promoting local agricultural economies. (Supplementary information).

Previous research on fruit quality testing has focused on both intrinsic and external characteristics. Intrinsic characterization indicators, such as nutrient content, moisture content, ripeness, drug residues, and internal lesions, are critical to assessing fruit quality². Techniques like Near Infrared Spectroscopy have been widely used to determine soluble solids content, acidity, sugar and moisture content in fruits^{3–5}. Other methods, such as Hyperspectral Imaging⁶, Mass Spectrometry^{7,8}, Liquid chromatography^{9,10}, and Gas Chromatography-Mass Spectrometry¹¹ have shown promise for assessing internal qualities and identifying organic compounds. In addition, methods such as Nuclear Magnetic Resonance¹², ultrasonic¹³, and X-ray¹⁴ can also be realized to accurately detect intrinsic characteristics of fruits. Although these methods can provide accurate and detailed measurements, they often involve high equipment costs, frequent calibration, and complex data analysis, which limits their practical application in large-scale fruit quality assessment.

In contrast, external quality inspection methods focus on visual characteristics, such as color, texture, size, and shape, as well as damage, decay, and disease, to evaluate fruit quality during harvesting and grading^{15,16}. However, manual visual inspection remains the most common method despite its high cost, low efficiency and susceptibility to human error^{17,18}.

Sweetness is an important factor in determining the quality of fruits and in addition to some of the previously mentioned methods in the literature, some new techniques, as, Fourier Transform Infrared Spectroscopy¹⁹, Enzyme Biosensor²⁰, Nuclear Magnetic Resonance²¹, Mathematical model construction^{22,23}, are also used

¹Zhejiang Provincial Key Laboratory for Cutting Tools, Research Institute of Motor and Intelligent Control Technology, Taizhou University, Jiaojiang, Taizhou 318000, Zhejiang, China. ²School of Light Industry, Harbin University of Commerce, Harbin 150028, China. ³WenZhou Polytechnic, WenZhou 325035, ZheJiang, China. ✉email: 2021000437@wzpt.edu.cn

for sweetness detection of fruits. In recent years, Computer/Machine Vision and Artificial Intelligence (AI) techniques have been extensively studied for crop quality assessment, Such as quality inspection of agricultural products^{24–26}, identification of crop diseases and pests^{27–30}, and so on. Meanwhile, research on these techniques for fruit brix detection has also progressed, Yu et al.³¹ developed a deep learning method combining Stacked Auto-Encoders and Fully Connected Neural Network (CNN) for predicting the soluble solids content of Kuril balsam pear. Liu et al.³² optimized a multivariate calibration model for non-destructive testing of avocado sugar content using synergy interval partial least squares-competitive adaptive reweighted sampling method. Munawar et al.³³ focused on the performance of three different regression methods, least squares regression model, Support Vector Machine Regression and Artificial Neural Network, for predicting the total acidity of mango was investigated. Nguyen et al.³⁴ implemented a predictive model for the total acidity of mango using a Random Forest classifier was implemented to accurately classify the sweetness of mango. In the above work, the researcher measured the soluble solids content and classified the samples into different sweetness levels based on their content, which in turn enables the correlation between the sugar content of the samples and the colour image data, which is a commonly used approach in the detection of image inspection of agricultural products³⁵.

To summarize, these advancements, however, have not yet been applied to Linhai mandarin oranges, and the existing models are often not tailored for specific fruit types or environmental conditions. Therefore, this paper aims to explore CNN specifically for predicting the sweetness of Linhai mandarin oranges. By collecting samples of Linhai mandarin oranges to construct a dataset that is images and sugar content related, this paper aims to propose a new CNN model and detection method that enables accurate, efficient, and cost-effective sweetness detection, achieving the possibilities for large-scale quality assessment of Linhai mandarin oranges.

The main objectives and contributions of this paper are as follows:

- (1). Taking the Linhai mandarin orange produced in Linhai City, a city on the southeast coast of China, as a case study, 1,000 Linhai mandarin orange were collected in the field to form a dataset with 5,022 images and classified into five categories according to the percentage of sugar concentration.
- (2). The Attention for Orange (AO) attention mechanism model is innovatively proposed to capture the details of orange features.
- (3). Based on the AO attention mechanism module, the Multiscale Feature Optimization (MFO) module is innovatively proposed, aiming to further enhance the capability of capturing information and details of oranges from the spatial dimension.
- (4). The proposed AO and MFO modules were applied to the improvement of the VGG13 to obtain the VGG-MFO-Orange network model, which resulted in a citrus sweetness prediction accuracy of 86.8%.

Related works

Convolutional neural networks

A CNN is a neural network architecture built for image processing, especially for recognition and classification tasks^{36–38}. CNNs are characterized by their ability to automatically learn discriminative, robust features directly from image data without the need for human-engineered feature extraction. CNNs have been shown to outperform traditional feature engineering methods in distinguishing plant species and diagnosing plant diseases^{39,40}. A typical CNN architecture consists of three layers: a convolutional layer, a pooling layer, and a classification layer⁴¹. Typically, CNNs have fewer parameters compared to traditional neural networks of the same size, which results in superior model performance.

Attention mechanisms

There are three main types of attention mechanisms: channel attention, spatial attention, and the fusion of these two forms. After an image is processed by CNN, important features are retained in each layer. Accurate feature parameter calculation is an important step in deep learning, and the attention mechanism can accurately weight the data features to help the model filter out key information from a large amount of data, thus improving performance. Currently commonly used attention mechanisms are:

- (a) Squeeze-and-excitation modules⁴², which optimize feature extraction by dynamically adjusting channel weights using global average pooling and two-layer fully-connected networks, thus solving the problem of uneven information weights between channels;
- (b) Combining channel attention and spatial attention through global average pooling and maximum pooling, as well as one-dimensional and two-dimensional convolutional operations, to dynamically adjust the channel and spatial dimensional feature weights, thus optimizing the feature expression of the convolutional block attention module⁴³.
- (c) Efficient channel attention modules, which efficiently implement the channel attention mechanism through one-dimensional convolution to reduce information loss and decrease the parameter scale, thus adaptively adjusting the feature mapping channel weights⁴⁴.

Feature extraction

Feature extraction is a key step in deep learning and image processing. It aims to capture and distill key pieces of information from the original data to construct a feature space that can effectively express the attributes of the data. In CNNs, feature extraction is mainly implemented through convolutional layers, which utilize multiple filters to capture different features in the image, such as edges, corner points, etc.; in addition, the spatial dimension of the feature map is reduced by a pooling layer to enhance the robustness of the features^{45,46}. Eventually, a fully connected layer integrates these features to provide output for classification or decision-making tasks.

VGGNet

VGGNet⁴⁷ is a convolutional neural network developed by the Visual Geometry Group at the University of Oxford. It employs a cascaded network structure by using small 3×3 convolutional filters and introducing a pooling layer after every 2 to 3 convolutional layers. VGGNet usually consists of 16 or more convolutional, pooling and fully connected layers, such as in the classical VGG16 and VGG19 models.

In this paper, advanced deep learning techniques to enhance the prediction of citrus sugar content is integrated. The attention mechanism improves feature selection by dynamically adjusting the importance of key features such as the peel's texture and color. Multiscale feature optimization (MFO) captures information at various scales, enabling the model to better understand and utilize the diverse features of the citrus peel. The VGG13 model, with its robust feature extraction capabilities and the benefit of transfer learning, serves as a strong foundation for these enhancements. By combining these methods, to create a model for achieving effectively captures and fuses key features to improve prediction accuracy and model performance.

Material and methods

Overview

To integrate a convolution neural network model with non-destructive testing technology for accurately predicting the sugar content of Linhai mandarin oranges, this paper begins with dataset collection, including images collection of Linhai mandarin oranges and its sugar content measurements. Feature extraction, model training, and validation are subsequently conducted with multiple network architectures, including both classical and the proposed model. This is followed by comprehensive model testing and performance comparison to evaluate the effectiveness of the proposed model against classical models. Finally, the sugar content prediction ability of the proposed model is detailed discussed from the perspective of convolution visualization. The flow diagram of this paper is shown as Fig. 1.

Data description

As a case study, 1000 Linhai mandarin oranges were randomly sampled from orchards in Xicen Village, Yanjiang Town, Linhai City, which is well-known as the main mandarin-producing area in the southeast coastal region of China. All the mandarin orange samples were then transported to the visual inspection laboratory of Taizhou College. Losses during transportation were removed, resulting in 990 samples, and then all the mandarin oranges were numbered sequentially. It should be noted that all data collection for these samples was completed within 24 h to avoid the influence of time variation on the experimental results.

The mandarin image acquisition process is shown in Fig. 2. Prior to data collection, the image acquisition system, shown in Fig. 2b, was started and run for one hour to ensure stability of the lighting and experimental equipment. Image acquisition was performed on all the mandarin oranges using a camera (MV-CE200-10UC, Hangzhou Hikvision Digital Technology Co., Ltd, 5184×3456 pixels). During the acquisition process, the mandarin oranges were manually placed under the camera one by one in random orientations with the umbilicus facing upwards to demonstrate the various possible poses when grading the fruits.

The sugar content testing of mandarin oranges was subsequently carried out. Specifically, five segments of each mandarin orange were randomly selected and placed into a mortar and pestle, and after grinding, filtering, and obtaining the juice, the refractive index of the juice containing soluble solids was measured using an Abbe refractometer (WYA-2WAJ, Shanghai Precision Scientific Instruments Co., Ltd, with sugar content detection range 0–95%, accuracy 0.25%) so as to indirectly obtain the percentage of sugar concentration of the Linhai mandarin oranges. Five replicate measurements were carried out on the juice of each sample collected. After removing the maximum and minimum values using the trimmed mean method, the average of the remaining determinations was calculated; the results obtained are shown in Table 1. The obtained image samples were

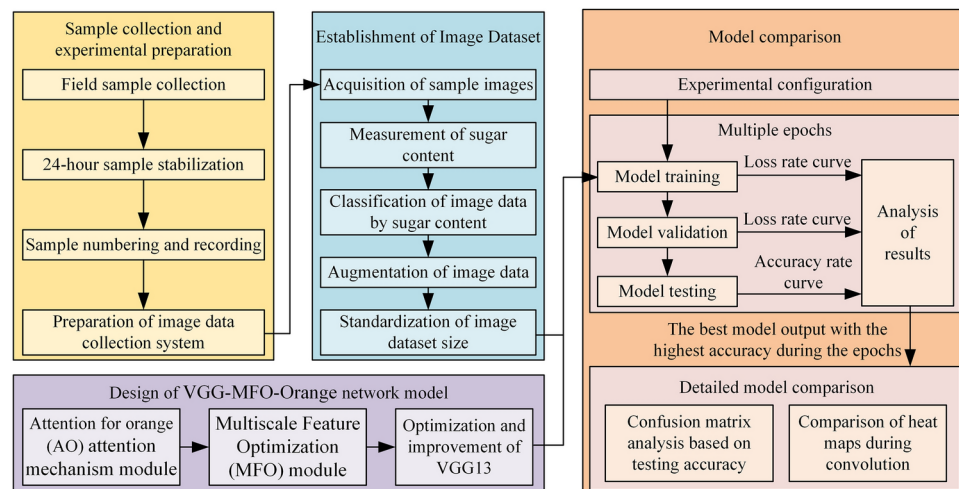


Fig. 1. Flow diagram of this paper.

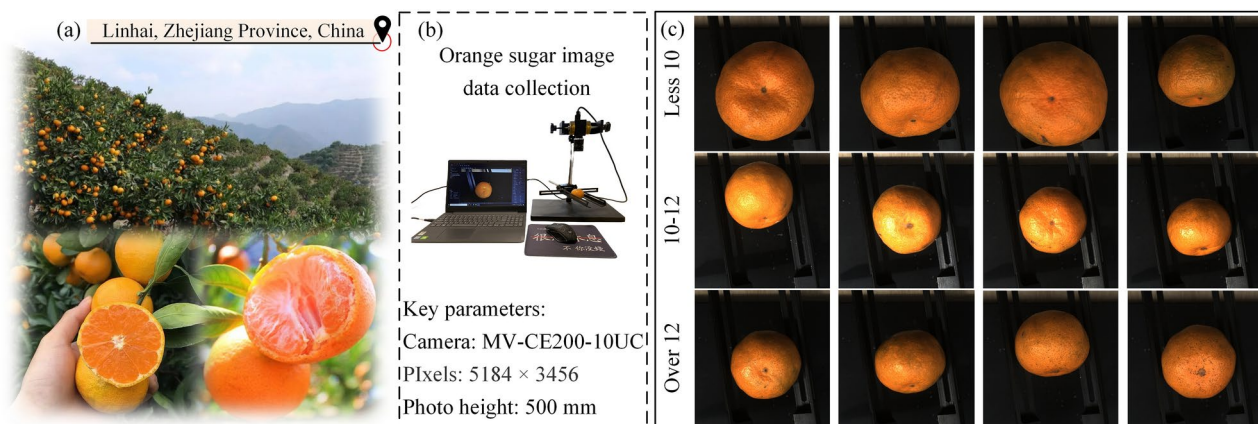


Fig. 2. Image data collection of Linhai mandarin oranges. (a) The well-known major production area of mandarin oranges in the southeastern coastal area of China. (b) Image acquisition system. (c) Subset of collected Linhai mandarin orange sample images.

Measured value range	≤9	9–9.5	9.5–10	10–10.5	10.5–11	11–11.5	11.5–12	12–12.5	12.5–13	>13
Sample size	6	15	54	148	235	274	164	66	20	8

Table 1. The detected sugar content data of the collected 990 Linhai mandarin oranges.

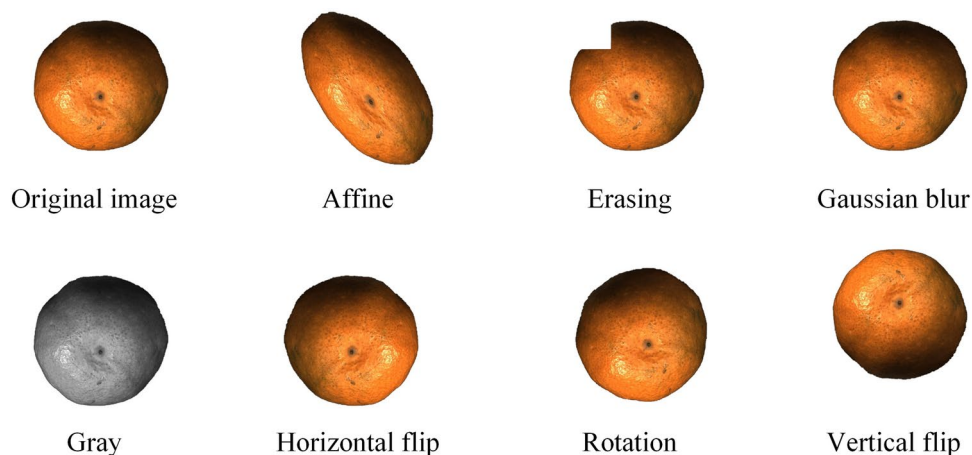


Fig. 3. Augmentation of the original Linhai mandarin orange image dataset.

classified into three categories according to the sugar levels “less 10” (signifying sugar content of less than 10%), “10–12” (sugar content of 10–12%), and “over 12” (sugar content of greater than 12%), as shown in Fig. 2c, which served as the basis for the classification of the quality detection of the Linhai mandarin oranges.

In order to ensure the diversity of sample images and to avoid overfitting, we adopted an offline image augmentation method involving affine transformation, erasure, Gaussian blur, grayscale, horizontal flipping, rotation, and vertical flipping. The original image dataset was augmented, with an example of this augmentation shown in Fig. 3. After manually filtering and removing the invalid augmented images, the final numbers of “less 10,” “10–12,” and “over 12” were 1500, 1642, and 1880, respectively. Finally, the pixel size of all the images was uniformly adjusted to 256 × 256 in preparation for CNN training.

Proposed model

Attention for orange (AO) model

In order to realize the CNN model for capturing the details of orange features, we first designed the Attention for Orange (AO) attention mechanism module. This mainly included Conv_block and LR_block. The specific design architecture is shown in Fig. 4, and the parameters are shown in the Table 2. Among them, Conv_block utilizes group convolution to process spatial information, and preserves local features and integrates channel

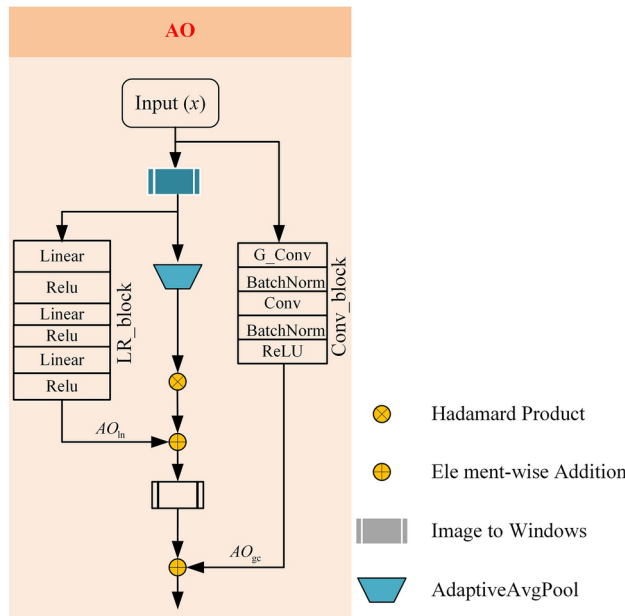


Fig. 4. Attention for orange module.

Name	Layer type	Input	Output				
LR_block	Linear	32	8				
	ReLU	-					
	Linear	8	4				
	ReLU	-					
	Linear	4	8				
	ReLU	-			Kernel	Stride	Padding
Conv_block	G_Conv	32	32	(3, 3)	2	1	32/64/128/256
	BatchNorm	32	32	-			
	Conv	32	32	(1, 1)	1	0	-
	BatchNorm	32	32	-			
	ReLU	-					

Table 2. Parameters of attention for orange (AO).

information through 3×3 convolution and 1×1 convolution, respectively. Meanwhile, LR_block uses fully connected layer degradation and ReLU to enhance nonlinear feature representation. Overall, the AO module combines the outputs of Conv_block and LR_block to fuse local and global features through element-level summation to improve computational efficiency and enhance detail attention.

In addition, to further enhance the ability to capture global information, the AO module was designed with a AdaptiveAvgPool mechanism that obtained the global average pooling result of the channel dimension and summed it with the output. The result was then transposed back to the original shape to sum with the output of the convolution, thus integrating global information.

A summary of the attention mechanism equation for the AO module is shown below:

$$AO(y) = AO_{gc}(x) + (AO_{ln}(x) + AO_{ln}x P_{Avg_pool}(x^T))^T \tag{1}$$

Among which x^T denotes permute, transpose, reconfigure, reshape; $AO_{gc}(\cdot)$ represents the Conv_block result; $AO_{ln}(\cdot)$ represents the LR_block result; and P_{Avg_pool} denotes adaptive pooling. $AO_{ln}(\cdot)$ can be expressed as:

$$AO_{ln}(y) = LR_3(LR_2(LR_1(x))) \tag{2}$$

where LR consists of Linear and ReLU. $LR_i(\cdot)$, $i \in \{1, 2, 3\}$.

Multiscale feature optimization (MFO) model

Based on the AO attention mechanism module, this work implemented a Multiscale Feature Optimization (MFO) module (Fig. 5). This involved segmenting and flipping the input feature maps, and then applying

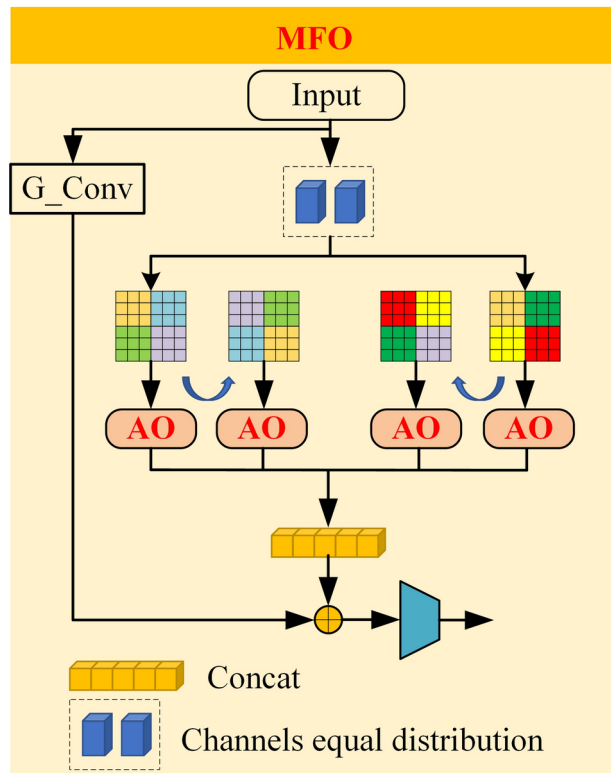


Fig. 5. Multiscale Feature Optimization module.

the AO module independently on the feature maps before and after flipping to enhance the capture of spatial information and object details. The processed feature maps were spliced in the channel dimension to recover the number of channels. Meanwhile, the residual information of the original feature maps was preserved by 1×1 convolution and summed with the spliced feature maps. The final global feature aggregation was performed by average pooling operation. This design not only enhanced the robustness of the model to spatial variation, but also further improved the recognition of orange features through residual concatenation, optimizing the feature extraction effect of the CNN.

Specifically, in MFO module, the input features are divided into two pieces after the “Channels equal distribution operation”, and then the left and right pieces are diagonally flipped to obtain a total of four parts to be processed, which are then inputted into four OA modules respectively. It is worth noting that the G_Convs Groups of these four OA modules are set as 32, 64, 128 and 256 respectively, as shown in Table 2.

The feature calculation in MFO can be expressed as:

$$P_{Avg_pool} \left(G_Conv (Feature_i) + concat \left(AO (Feature_i^1), AO (Feature_i^2), AO (Feature_i^3), AO (Feature_i^4) \right) \right) \quad (3)$$

where G_Conv is the group convolution residual. Its Groups take the same value as that of Groups in AO, and can be expressed as:

$$Groups_{MFO} (j) = Groups_{AO} (j) = \begin{cases} 32 & \text{if } j=1 \\ 64 & \text{if } j=2 \\ 128 & \text{if } j=3 \\ 256 & \text{if } j=4 \end{cases} \quad (4)$$

VGG-MFO-orange model

The VGG-MFO-Orange model proposed in this paper is an innovative improvement involving the addition of our designed AO and MFO modules to the classic VGG13 architecture. While maintaining the original convolutional layer structure of VGG13, the model introduces Batch Normalization (BN) to enhance the generalization ability and training stability of the model. The network structure of the proposed model is shown in Fig. 6, and the network parameters are shown in Table 3.

Overall, four MFO modules were integrated into the model. These modules act on different layers of feature maps, which in turn enhance the extraction capability of recognizing key feature information. Meanwhile, in order to enrich the classification strategy, this study also introduced a new classifier, FC_new, which was

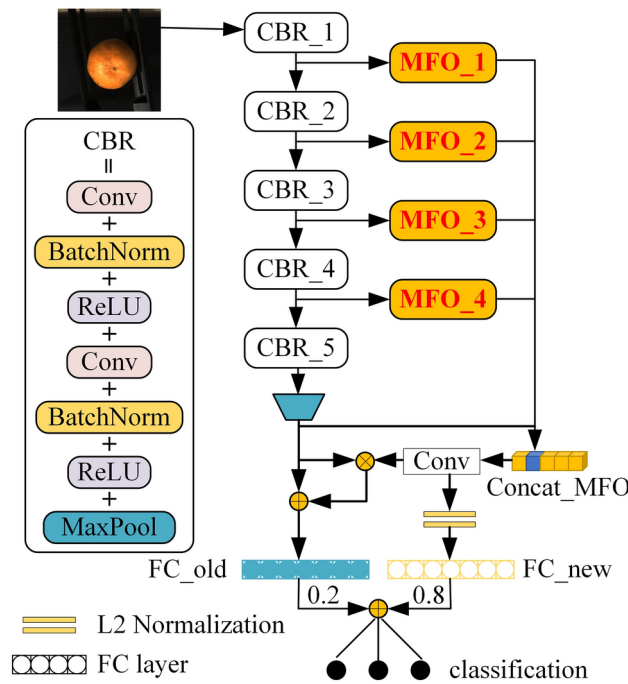


Fig. 6. Network structure diagram of the proposed model.

combined with the existing FC_old, enabling the model to employ both traditional classification methods and cosine similarity-based classification methods. This comprehensive design strategy not only fully utilizes the depth and complexity of the VGG architecture, but also increases its competitiveness in image classification tasks through the introduced AOs, MFOs, and classifiers.

The flow of the model generated in this paper is shown below:

$$Merge = w \left(concat \left(P_{Avg_pool1}, P_{Avg_pool2}, P_{Avg_pool3}, P_{Avg_pool4}, P_{Avg} \right) \right) \tag{5}$$

The attentively weighted feature maps conv1_AO, conv2_AO, conv3_AO, and conv4_AO, and the globally averaged representation of the feature map averages, are spliced along the channel dimensions to obtain the fused feature map. The fused feature maps are then passed to the final convolutional layer, whose dimension is reduced from 992 to 512 to obtain the output feature map Merge:

$$Weighted_Merge = P_{Avg} + P_{Avg} Merge \tag{6}$$

$$logist_1 = FC(Weighted_Merge) \tag{7}$$

Through Eqs. (6) and (7), prediction result 1 is computed. Merge is operated with the earlier-obtained average, and the result is saved as a variable average; that is, average = average + average * Merge. Then, the result average is spread as a one-dimensional vector and passed to the classifier for prediction to obtain result 1.

$$FC1 = \sigma(FC(\cdot)_{Weight}) \tag{8}$$

$$logist_2 = FC1(\sigma(Merge)), \sigma(\cdot) \rightarrow Normalize \tag{9}$$

Through Eqs. (8) and (9), prediction result 2 is computed. The L2-paradigm normalization is directly performed on the feature map Merge, and then the normalized feature vectors are passed to the linear classifier to obtain the classification prediction result 2.

Finally, the cross-entropy loss of the computed result 1 with label is computed using Eq. (10) to obtain FC_new. The cross-entropy loss of the computed result 2 with label is computed using Eq. (11) to obtain FC_old:

$$FC_{old} = L_{CE}(logist_1, label) \tag{10}$$

$$FC_{new} = L_{CE}(logist_2, label) \tag{11}$$

where $L_{CE}()$ is the cross entropy. The proportion of FC_{new} and FC_{old} in the final categorization result is represented by Eq. (12):

$$FC_{total} = 0.8 \times FC_{new} + 0.2 \times FC_{old} \tag{12}$$

Block name	Layer	Input	Output	Kernel	Stride	Padding
CBR_1	Conv 1_1	3	64	(3, 3)	1	1
	BatchNorm 1_1	64	64	-	-	-
	ReLU 1_1	-				
	Conv 1_2	64	64	(3, 3)	1	1
	BatchNorm 1_2	64	64	-	-	-
	ReLU 1_2	-				
	MaxPool 1	-	-	(2, 2)	2	0
CBR_2	Conv 2_1	64	128	(3, 3)	1	1
	BatchNorm 2_1	128	128	-	-	-
	ReLU 2_1	-				
	Conv 2_2	128	128	(3, 3)	1	1
	BatchNorm 2_2	128	128	-	-	-
	ReLU 2_2	-				
	MaxPool 2	-	-	(2, 2)	2	0
CBR_3	Conv 3_1	128	256	(3, 3)	1	1
	BatchNorm 3_1	256	256	-	-	-
	ReLU 3_1	-				
	Conv 3_2	256	256	(3, 3)	1	1
	BatchNorm 3_2	256	256	-	-	-
	ReLU 3_2	-				
	MaxPool 3	-	-	(2, 2)	2	0
CBR_4	Conv 4_1	256	512	(3, 3)	1	1
	BatchNorm 4_1	512	512	-	-	-
	ReLU 4_1	-				
	Conv 4_2	512	512	(3, 3)	1	1
	BatchNorm 4_2	512	512	-	-	-
	ReLU 4_2	-				
	MaxPool 4	-	-	(2, 2)	2	0
CBR_5	Conv 5_1	512	512	(3, 3)	1	1
	BatchNorm 5_1	512	512	-	-	-
	ReLU 5_1	-				
	Conv 5_2	512	512	(3, 3)	1	1
	BatchNorm 5_2	512	512	-	-	-
	ReLU 5_2	-				
	MaxPool 5	-	-	(2, 2)	2	0
Avg	AdaptiveAvgPool	-				
Classifier (FC1)	Linear	512	3	-	-	-
Classifier (FC2)	Linear	512	3	-	-	-

Table 3. Parameters of the VGG-MFO-orange network.

Experimental results and analysis

Experimental configurations

The model described in this paper was trained and validated in the environment shown in Table 4, and the experimental parameters of the network model were set as shown in Table 5.

Experiments on the Linhai mandarin orange image dataset

Based on the work in Sect. 3.2, the expanded Linhai mandarin orange image dataset was divided into a training set and a validation set in a 7:3 ratio. To further validate the effectiveness of the proposed method, six influential CNNs, namely DenseNet121, MobileNet_v2, ResNet50, ShuffleNet, VGG13, and VGG13_bn, were considered in comparison experiments.

The training and validation loss curves during the experiment are shown in Figs. 7 and 8, in which a significant difference in the performance of these models within 100 epochs can be seen. On the training loss curves, the loss values of all models decreased rapidly, indicating that these models fit better on the training dataset. However, as can be seen from the validation loss curve, the validation loss did not decrease as steadily as the training loss. There were even marked fluctuations and rises in some models (e.g., DenseNet121, ResNet50, and VGG13), suggesting that there may have been overfitting in them. The model presented in this paper exhibited a rapid decline and smooth final loss values in both the training and validation loss curves, indicating that the network, with good learning ability, could better adapt to our dataset. In contrast, the validation loss curves of

Environment	Configuration
Programming language	Python 3.10.9
Operating system	Windows 10 Professional
Framework	Torch: 1.21.1 + cu113
GPU	NVIDIA GeForce RTX4070Ti
CPU	13th Gen Intel(R) Core(TM) i9-13900 K
Memory	32 GB DDR5
Library called	Torchvision: 0.13.1 + cu113; NumPy: 1.23.5

Table 4. The experimental environment configuration.

Training parameters	Details
Iteration	100
Input image size (pixels)	256 × 256
Learning rate	0.000005
Optimizer	Adam
Classification number	3
Batch size	48
Epoch	100

Table 5. The main parameter settings of the proposed model.

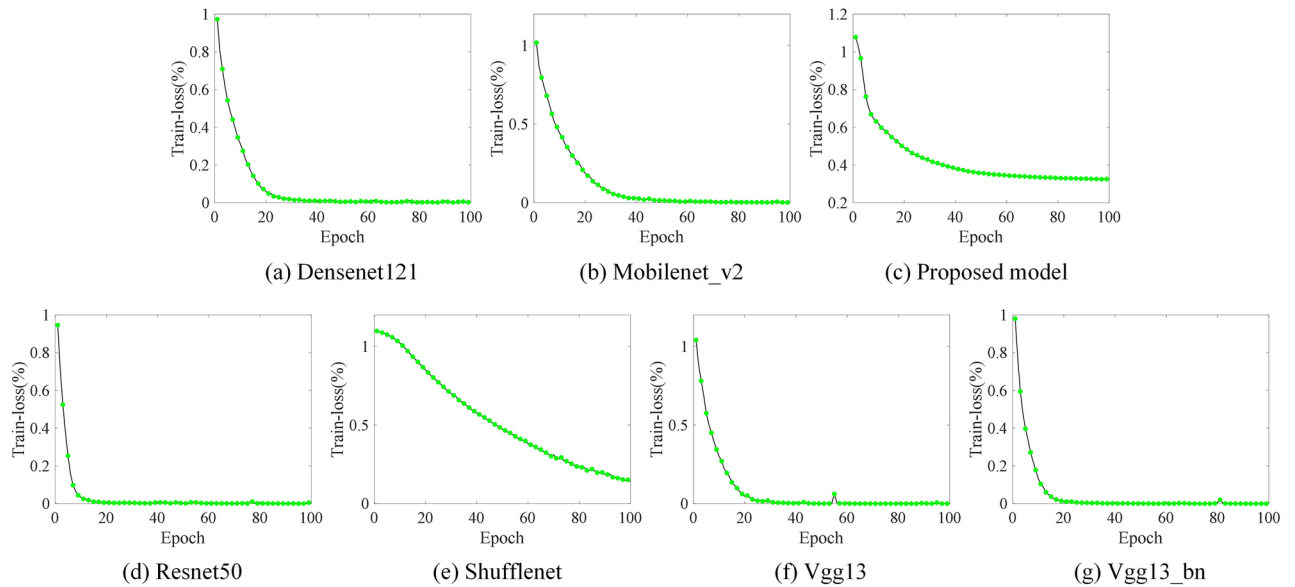


Fig. 7. Loss rate curves for each network during the training process.

VGG13 and VGG13_bn fluctuated more, indicating that these models had insufficient generalization ability for the validation dataset and may have suffered from overfitting or unstable optimization. Finally, ShuffleNet and MobileNet_v2 did not show a significant decrease in validation loss despite their good performance in the training phase, suggesting that these lightweight models may have failed to adequately learn the complex features of the data in the given task.

The accuracy curve produced during network validation is shown in Fig. 9, in which it can be seen that the network designed in this paper performed well on the test set. Its accuracy increased steadily with the increase in training cycles, finally reaching a relatively high level; this shows that the proposed network had good generalization ability on the data of mandarin orange sugar classification images. Meanwhile, DenseNet121 and ResNet50 also showed good performance: Their accuracies gradually increased with training and, although the final accuracies were slightly lower than those of our model, they still showed a strong learning ability. MobileNet_v2 showed a faster increase in accuracy at the beginning of the training period, but then the growth slowed down, and the final accuracies were slightly lower than those of DenseNet121 and ResNet50; this may

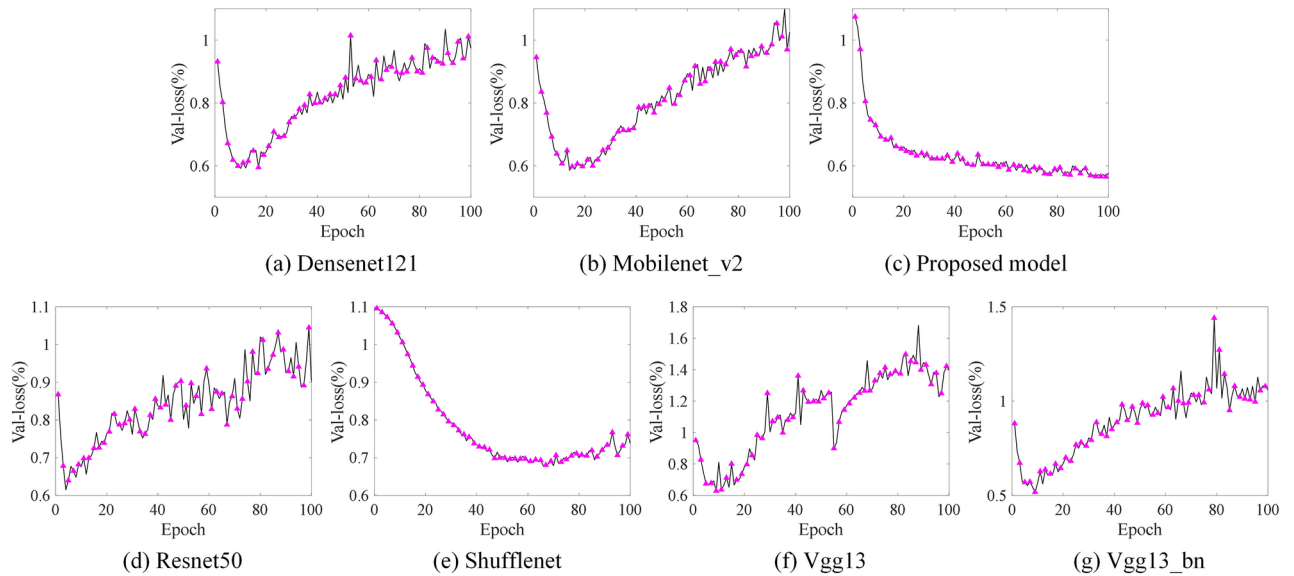


Fig. 8. Loss rate curves for each network during the validation process.

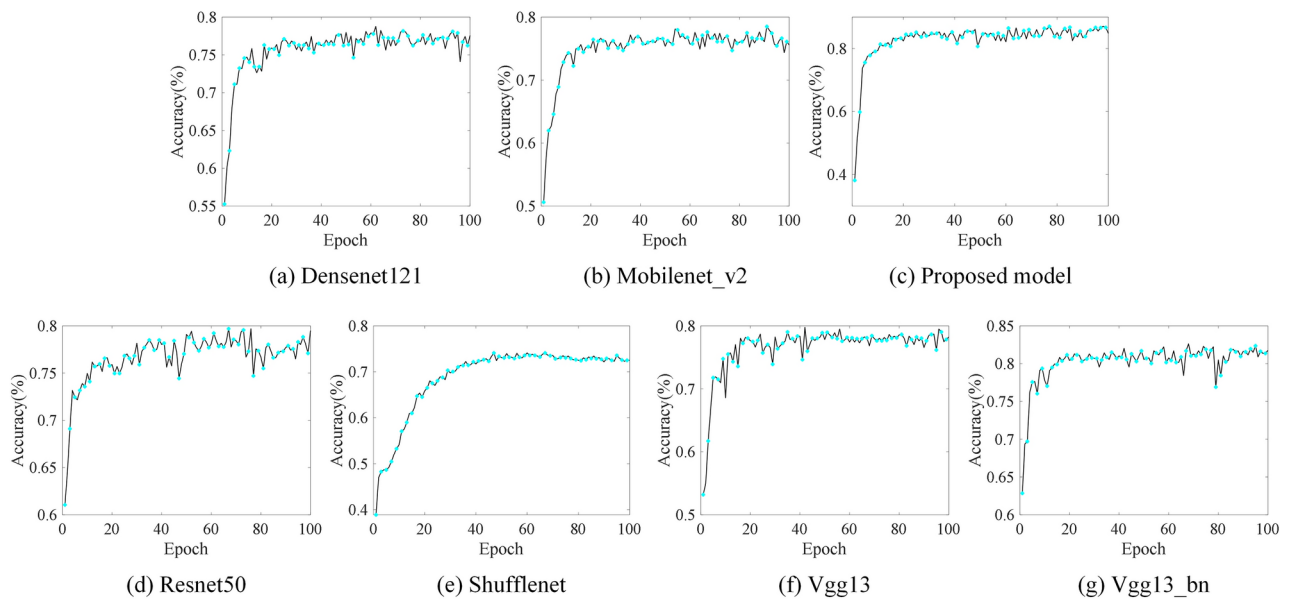


Fig. 9. Accuracy curves for each network during validation.

imply that the network is limited in some aspects. ShuffleNet’s accuracy growth was slower, and its final accuracy was relatively low, which may indicate that the network did not perform as well as the other networks in dealing with this task. The accuracy growth of VGG13 and its variant VGG13_bn was also slower, and their final accuracies were not very high, which may be related to their network structure.

During the model validation process, the settings were set to periodically evaluate model performance for each epoch. We selected the output and saved the model that achieved the highest accuracy. The confusion matrix of the validation results plotted by extracting the output file is shown in Fig. 10. From the results, the model proposed in this paper correctly classified 483 samples in the category “over 12,” which was second only to ResNet50. Its accuracy in the category “10–12” was also very high, correctly classifying 479 samples, which was the highest among all networks. The proposed model also had a low misclassification error rate; specifically, there was only one case where the category “10–12” was incorrectly identified as “over 12.” Among the other networks, DenseNet121 and MobileNet_v2 performed relatively well. However, they were slightly worse in the “over 12” category, correctly classifying 461 and 470 samples, respectively. In contrast, VGG13_bn and ResNet50 performed poorly on the “less 10” category, with high misclassification rates.

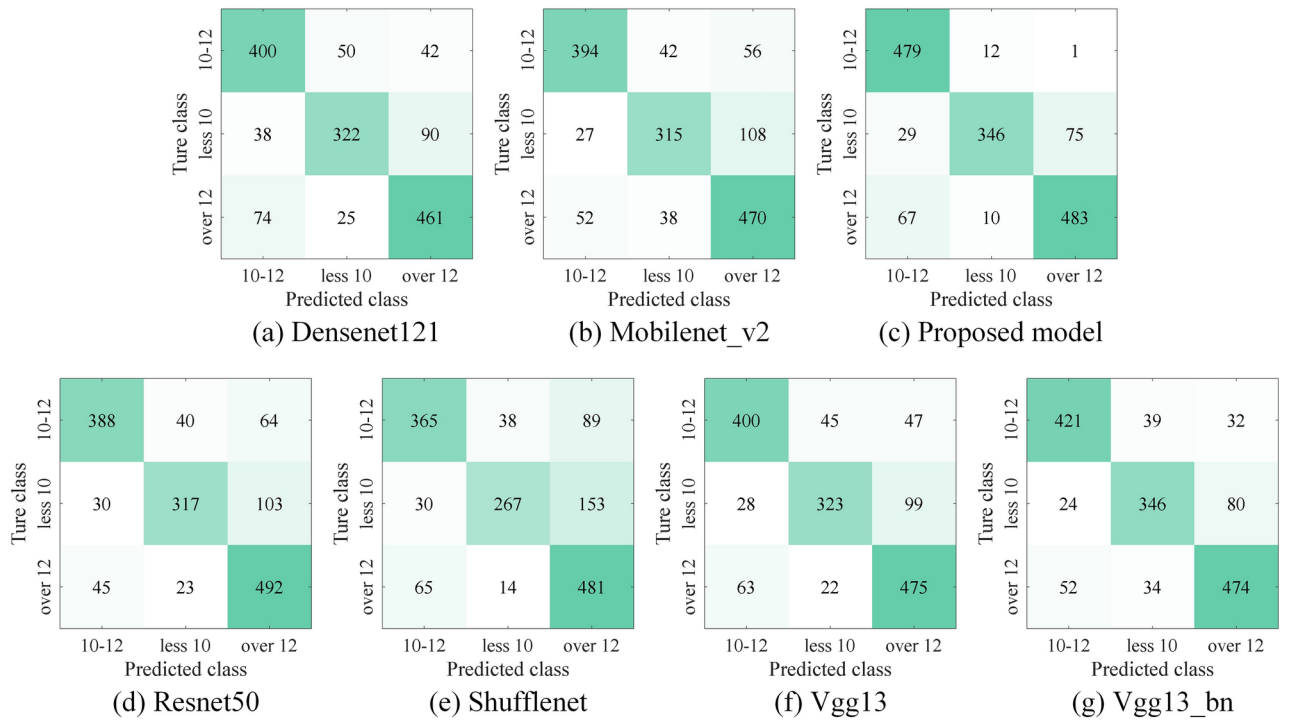


Fig. 10. Confusion matrix of test results where each network achieved the highest accuracy.

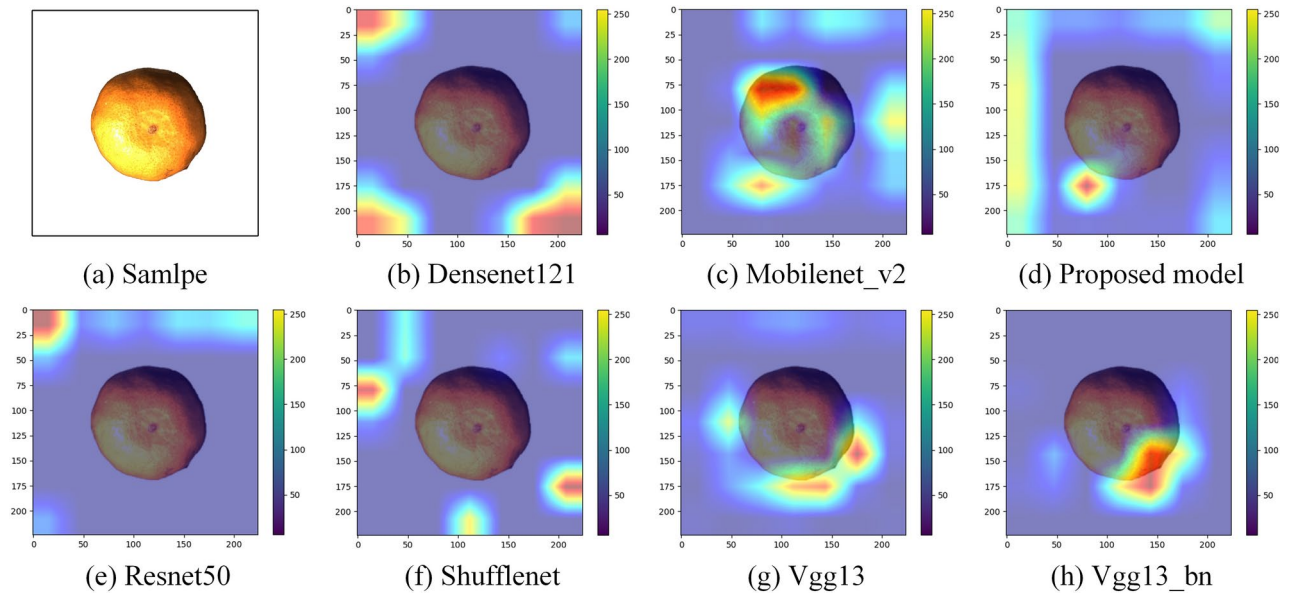


Fig. 11. Comparison of final convolutional heat maps of each network.

In order to compare the feature extraction ability of different networks, we chose to use the heat map of the image in the final convolution step of each network for horizontal comparison. A sample from the test set “10–12” was randomly selected as the uniform input, and the heat map (All the heat maps are created by PyCharm Community Edition, 2023.1, <https://www.jetbrains.com/pycharm/download>) of the final convolution output of each network is shown in Fig. 11. From the results of the comparison, the proposed model showed a more focused area of attention in feature extraction, especially in the edge and bottom regions of the fruit, which indicates a stronger feature extraction ability. In contrast, DenseNet121 and MobileNet_v2 had more pronounced attention toward the top and edge regions of the fruit, although their focus was more dispersed. The heat maps of ResNet50 and ShuffleNet, on the other hand, show a more even distribution of attention, but are not

as prominent as the proposed model in key feature areas. The heat maps of VGG13 and VGG13_bn heat maps appear to have a more dispersed focus of attention, with some areas of attention failing to focus even on the fruit.

The performance of the network model proposed in this paper, as well as the performance of each of its modules, was compared and analyzed vertically. Taking a random subset of data in the training and validation dataset as an example, heat maps of the last layer convolution results of CBR_1, CBR_2, CBR_3, CBR_4, and CBR_5 were output; the results are shown in Fig. 12. From the vertical comparison results, the ability of our network model to extract features gradually in different convolutional layers was clearly demonstrated. In the initial CBR_1 and CBR_2, the model mainly focused on the overall contour and basic shape features of the fruit, and the feature extraction was more uniform and did not highlight the key regions. As the convolutional layers deepened, from CBR_3 to CBR_5, the network gradually focused on more detailed features. In particular, in CBR_4 and CBR_5, the heat maps show that the model had a more pronounced focus on the edges of the fruits and key feature regions (e.g., surface depressions or color variations), indicating that the deeper convolutional layers were able to effectively capture complex image features. This gradual deepening of feature extraction, especially the consistency over the validation set data, demonstrated the robustness and stability of the model over different datasets. Overall, the feature extraction process of the model in each convolutional layer exhibited an effective feature extraction mechanism from shallow to deep and from whole to local, which verified the rationality and effectiveness of its design.

In order to evaluate the improvement effect of the added MFO module on VGG, the heat map results of each layer of MFO were output, as shown in Fig. 13. In the MFO_1 results, it can be seen that the MFO module quickly extracted orange peel feature, especially for sample1-2 and sample1-1, while as a comparison, the CBR_1 results in Fig. 12 did not extract too much feature. In the MFO_2 results, it can be seen that all samples have extracted key feature, which performs well compared to the CBR_2 results, indicating the success of the MFO module. Subsequently, the performance of MFO_3 can be thought as normal, while then, CBR_4 performed better than MFO_4. Finally, according to Eq. (5), the Concat_MFO layer outputted the concatenated results of MFO_1, MFO_2, MFO_3, MFO_4, and CBR_5. It can be seen that compared with the results of CBR_5, the result obtained

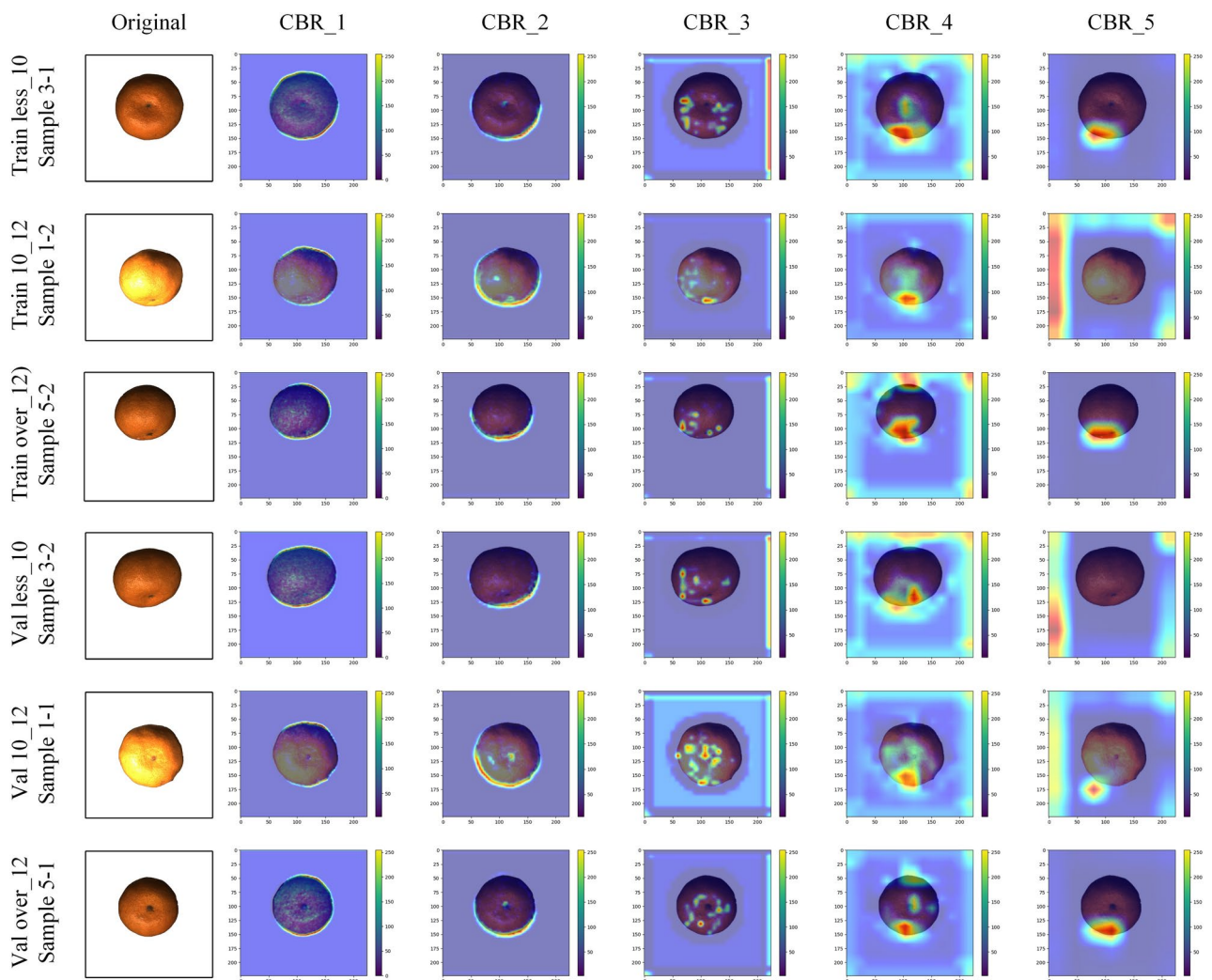


Fig. 12. Heat map results for each CBR result of the designed network.

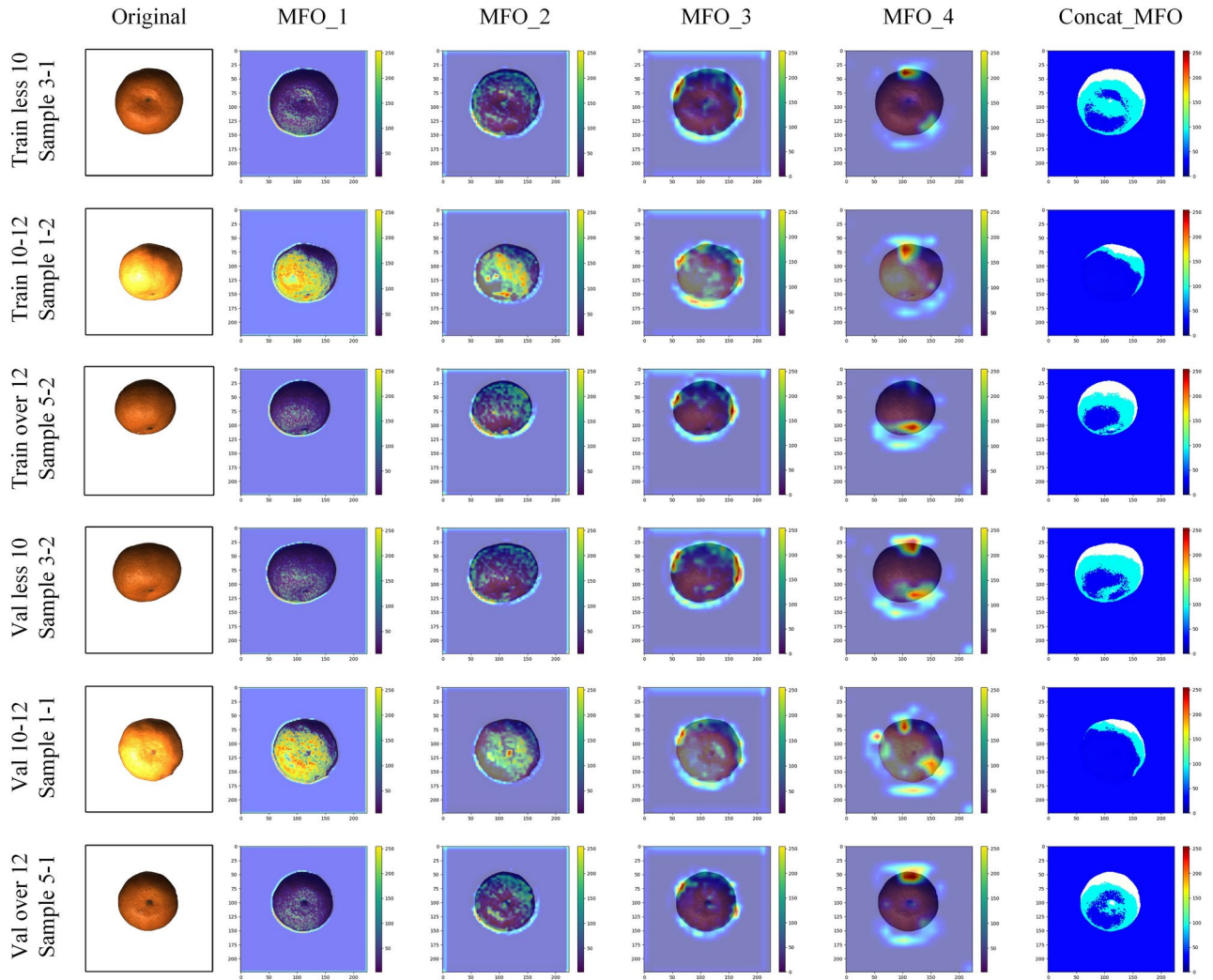


Fig. 13. Heat map results for each MFO result of the designed network.

by this operation can make the key features extracted fall around the orange peel instead of scattered like CBR_5. Overall, the proposed AO-MFO-VGG model is better than the original VGG in classification for Linhai mandarin oranges dataset.

In deep learning, precision, recall, and F1-score are important metrics for evaluating the performance of a model, and are obtained using the equations shown below:

$$\text{Precision}_i = \frac{TP_i}{TP_i + FP_i} \tag{13}$$

$$\text{Recall}_i = \frac{TP_i}{TP_i + FN_i} \tag{14}$$

$$\text{F1_score}_i = \frac{2 \times \text{Precision}_i \times \text{Recall}_i}{\text{Precision}_i + \text{Recall}_i} \tag{15}$$

where TP is True Positive, FP is False Positive, TN is True Negative, FN is False Negative, and *i* represents the first classification. In order to fully evaluate the performance of each model, the precision, recall, and F1-scores were further recalculated from the concepts of micro average, macro average, and weighted average as per Eqs. (16–24):

$$\text{Precision}_{\text{micro}} = \frac{\sum_{i=1}^n TP_i}{\sum_{i=1}^n (TP_i + FP_i)} \tag{16}$$

$$\text{Recall}_{\text{micro}} = \frac{\sum_{i=1}^n TP_i}{\sum_{i=1}^n (TP_i + FN_i)} \tag{17}$$

$$F1_{\text{micro}} = \frac{2 \times \text{Precision}_{\text{micro}} \times \text{Recall}_{\text{micro}}}{\text{Precision}_{\text{micro}} + \text{Recall}_{\text{micro}}} \quad (18)$$

$$\text{Precision}_{\text{macro}} = \frac{1}{n} \sum_{i=1}^n \text{Precision}_i \quad (19)$$

$$\text{Recall}_{\text{macro}} = \frac{1}{n} \sum_{i=1}^n \text{Recall}_i \quad (20)$$

$$F1_{\text{macro}} = \frac{1}{n} \sum_{i=1}^n F1 - \text{Score}_i \quad (21)$$

$$\text{Precision}_{\text{weighted}} = \sum_{i=1}^n w_i \times \text{Precision}_i \quad (22)$$

$$\text{Recall}_{\text{weighted}} = \sum_{i=1}^n w_i \times \text{Recall}_i \quad (23)$$

$$F1_{\text{weighted}} = \sum_{i=1}^n w_i \times F1 - \text{Score}_i \quad (24)$$

The micro-method calculates performance metrics from an overall perspective, it first accumulates the TP, FP, and FN across all categories, and then calculates precision, recall, and F1 scores based on these overall values. The macro-method calculates performance metrics from an individual perspective, it calculates precision, recall, and F1 scores separately for each category and then averages these scores. The weighted method calculates performance metrics from a weighted perspective, it calculates weighted precision, recall, and F1 scores based on the number of samples in each category.

The results calculated from Eqs. (16–24), shown in Table 6, demonstrate that the proposed model outperformed the other models in all indicators, especially in key evaluation indicators such as F1-score, precision, and recall. The F1-scores of the micro average, macro average, and weighted average of this model were 0.8708, 0.8690, and 0.8694, respectively, which were all higher than those of DenseNet121, MobileNet_v2, ResNet50, VGG13, and its with-batch normalized version, VGG13_bn. Specifically, the F1-scores of the proposed model in terms of its precision were 0.8791 and 0.8767 for the macro average and weighted average, respectively, further indicating its stability and accuracy in prediction. Meanwhile, the model also outperformed in metrics such as recall and F1-score, indicating that it has better overall performance in handling different categories of data.

The results of the model's number of parameters, final accuracy, and the time taken to train and validate an epoch for each network are shown in Table 7. It can be seen that the proposed model achieved relatively high accuracy with a relatively small number of parameters, with values of 10.04 M and 0.8708, respectively. However, its training time per epoch was longer, at 38 s, and its validation time was similar to the other models, at about 14 s. In contrast, ResNet50 had a final accuracy of 0.7969, a training time of 25 s, and a more moderate number of parameters. MobileNet_v2 and ShuffleNet had lower accuracies than the other models, at 0.7850 and 0.7976, respectively, despite having a smaller number of parameters and a shorter training time (both 20 s). VGG13 and its strip batch VGG13_bn had accuracies of 0.7876 and 0.8262, respectively; although these accuracies were higher than those of some of the other models, their training times were also relatively longer. Overall, the proposed model performed well in terms of accuracy, but required more training time, whereas ResNet50 struck a better balance between accuracy and training time.

Overall, the proposed model excelled in all metrics, especially outperforming other models in terms of accuracy, F1-score, and precision. It converged quickly on training and validation loss, and had strong

	DenseNet121	MobileNet_v2	Proposed model	ResNet50	VGG13	VGG13_bn
F1_Score _{micro}	0.7876	0.7850	0.8708	0.7969	0.7976	0.8262
F1_Score _{macro}	0.7856	0.7831	0.8690	0.7945	0.7958	0.8250
F1_Score _{weighted}	0.7869	0.7844	0.8694	0.7960	0.7969	0.8259
Recall _{micro}	0.7876	0.7850	0.8708	0.7969	0.7976	0.8262
Recall _{macro}	0.7839	0.7800	0.8683	0.7905	0.7930	0.8237
Recall _{weighted}	0.7876	0.7850	0.8708	0.7969	0.7976	0.8262
Precision _{micro}	0.7876	0.7850	0.8708	0.7969	0.7976	0.8262
Precision _{macro}	0.7899	0.7906	0.8791	0.8063	0.8026	0.8272
Precision _{weighted}	0.7888	0.7882	0.8767	0.8028	0.8002	0.8265

Table 6. Accuracy, precision, recall, and F1-score of each network.

Method	Parameters [M]	Accuracy	Training time per epoch [s]	Validation time per epoch [s]
DenseNet121	6.96	0.7876	26	14
MobileNet_v2	2.23	0.7850	20	13
Proposed model	10.04	0.8708	38	14
ResNet50	23.51	0.7969	25	13
ShuffleNet	5.35	0.7976	20	14
VGG13	128.96	0.7876	33	14
VGG13_bn	128.97	0.8262	36	14

Table 7. Results of the number of model parameters, final accuracy, and time spent training and validating an epoch for each network.

generalization ability, especially in the classification of key categories. Despite its long training time, the proposed model possessed excellent feature extraction ability that could effectively capture the edges and key features of fruits, demonstrating good learning ability and robustness.

Conclusions

With the increasing consumer demand for orange quality, the detection and grading of orange internal quality is becoming increasingly necessary. Deep learning techniques, especially convolutional neural networks (CNNs), have shown significant advantages in solving complex classification problems. In this paper, with the goal of enhancing the learning ability for citrus epidermal features, we investigated the migration learning of deep CNNs, and propose an improved “VGG-MFO-Orange” CNN model specifically for the non-destructive image recognition and sugar detection of Linhai mandarin orange. The specific contributions of this paper are as follows:

- A dataset is proposed. The dataset consists of two parts: orange appearance photos and orange fruit sugar content. The sugar content of the orange is predicted by the visual texture characteristics of the orange;
- An AO attention mechanism module for extracting the features of orange visual texture is proposed. The module divides the input into three sequential parts: the Linear Group module, Adaptive Mean Pooling module, and Convolutional Group module. The feature data of orange visual texture is then obtained through a series of operations such as Hadamard product and element-wise addition;
- An MFO module that divides the input data into left and right parts. The data from the left and right parts are operated by roll, after which the feature information is extracted through the attention mechanism. This is followed by generating the attention mapping with the output of each convolution block and the corresponding attention module to improve prediction accuracy.

The experimental results show that the model performed well on both of our own image datasets, with an accuracy of 0.8708. In the future, the proposed model will hopefully be deployed on mobile devices for a wider range of non-destructive fruit detection and classification tasks. This will help fruit farmers and distributors to optimize fruit grading and sales channels, as well as promoting the intelligent development of agricultural production.

In this paper, 1000 Linhai mandarin oranges were collected as samples from the same one village, which is limited for the vast Linhai mandarin orange production area, and the number of samples is not considered to be as a large scale. In the future works, the sample collection area will be expanded, the number of samples collected will be greatly increased, which will be used to further improve the accuracy of the proposed model, and the actual prototype detection device will be made to deploy the proposed model for practice.

Data availability

Data is provided within the manuscript or supplementary information files.

Received: 11 November 2024; Accepted: 27 March 2025

Published online: 06 April 2025

References

- Taizhou agricultural products win national recognition. http://taizhou.chinadaily.com.cn/2020-06/03/c_497419.htm.
- Zhang, B. et al. Principles, developments and applications of computer vision for external quality inspection of fruits and vegetables: A review. *Food Res. Int.* **62**, 326–343 (2014).
- An, J. et al. Prediction of sugar content of fresh peaches based on LDBN model using NIR spectroscopy. *J. Food Meas. Charact.* **18**, 2731–2743 (2024).
- Gomez, A. H., He, Y. & Pereira, A. G. Non-destructive measurement of acidity, soluble solids and firmness of Satsuma mandarin using Vis/NIR-spectroscopy techniques. *J. Food Eng.* **77**, 313–319 (2006).
- Wang, A. & Xie, L. Technology using near infrared spectroscopic and multivariate analysis to determine the soluble solids content of citrus fruit. *J. Food Eng.* **143**, 17–24 (2014).
- Xuan, G., Gao, C. & Shao, Y. Spectral and image analysis of hyperspectral data for internal and external quality assessment of peach fruit. *Spectrochim. Acta. A. Mol. Biomol. Spectrosc.* **272**, 121016 (2022).
- Capriotti, A. L., Cavaliere, C., Foglia, P., Piovesana, S. & Ventura, S. Chromatographic methods coupled to mass spectrometry detection for the determination of phenolic acids in plants and fruits. *J. Liq. Chromatogr. Relat. Technol.* **38**, 353–370 (2015).

8. Li, Y., He, Z., Zou, P., Ning, Y. & Zhu, X. Determination of seventeen sugars and sugar alcohols in fruit juice samples using hydrophilic interaction liquid chromatography-tandem mass spectrometry combining response surface methodology design. *Microchem. J.* **193**, 109136 (2023).
9. Chinnici, F., Spinabelli, U., Riponi, C. & Amati, A. Optimization of the determination of organic acids and sugars in fruit juices by ion-exclusion liquid chromatography. *J. Food Compos. Anal.* **18**, 121–130 (2005).
10. De Barros-Santos, R. G. et al. Ultra-fast determination of free carotenoids in fruit juices by rapid resolution liquid chromatography (RRLC): Method validation and characterization of Brazilian whole fruit juices. *Food Anal. Methods* **16**, 808–818 (2023).
11. Kang, H. S., Kim, M. & Kim, E. J. High-throughput simultaneous analysis of multiple pesticides in grain, fruit, and vegetables by GC-MS/MS. *Food Addit. Contam. Part A* **37**, 963–972 (2020).
12. Hou, J. et al. Compression damage mechanism and damage detection of *Aronia melanocarpa* based on nuclear magnetic resonance tests. *J. Food Meas. Charact.* <https://doi.org/10.1007/s11694-023-02213-y> (2023).
13. Yildiz, F., Uluisik, S., Özdemir, A. T. & İmamoğlu, H. Non-destructive testing (NDT): Development of a custom designed ultrasonic system for fruit quality evaluation. In *Nondestructive Quality Assessment Techniques for Fresh Fruits and Vegetables* (eds Pathare, P. B. & Rahman, M. S.) 281–300 (Springer Nature, 2022). https://doi.org/10.1007/978-981-19-5422-1_12.
14. Matsui, T., Kamata, T., Koseki, S. & Koyama, K. Development of automatic detection model for stem-end rots of ‘Hass’ avocado fruit using X-ray imaging and image processing. *Postharvest Biol. Technol.* **192**, 111996 (2022).
15. Ismail, N. & Malik, O. A. Real-time visual inspection system for grading fruits using computer vision and deep learning techniques. *Inf. Process. Agric.* **9**, 24–37 (2022).
16. Lorente, D., Escandell-Montero, P., Cubero, S., Gómez-Sanchis, J. & Blasco, J. Visible-NIR reflectance spectroscopy and manifold learning methods applied to the detection of fungal infections on citrus fruit. *J. Food Eng.* **163**, 17–24 (2015).
17. Arah, I. K., Ahorbo, G. K., Anku, E. K., Kumah, E. K. & Amaglo, H. Postharvest handling practices and treatment methods for tomato handlers in developing countries: A mini review. *Adv. Agric.* **2016**, 1–8 (2016).
18. Wang, Z., Jin, L., Wang, S. & Xu, H. Apple stem/calyx real-time recognition using YOLO-v5 algorithm for fruit automatic loading system. *Postharvest Biol. Technol.* **185**, 111808 (2022).
19. Song, S. Y., Lee, Y. K. & Kim, I.-J. Sugar and acid content of Citrus prediction modeling using FT-IR fingerprinting in combination with multivariate statistical analysis. *Food Chem.* **190**, 1027–1032 (2016).
20. Umar, L. et al. Amperometric microbial biosensor for sugars and sweetener classification using principal component analysis in beverages. *J. Food Sci. Technol.* **60**, 382–392 (2023).
21. Fakhari, H. I. et al. Universal 1H spin-lattice NMR relaxation features of sugar—a step towards quality markers. *Molecules* **29**, 2422 (2024).
22. Zhang, Y., Chen, Y., Wu, Y. & Cui, C. Accurate and nondestructive detection of apple brix and acidity based on visible and near-infrared spectroscopy. *Appl. Opt.* **60**, 4021 (2021).
23. Zhu, G. & Tian, C. Determining sugar content and firmness of ‘Fuji’ apples by using portable near-infrared spectrometer and diffuse transmittance spectroscopy. *J. Food Process Eng.* **41**, e12810 (2018).
24. Chaudhari, D. & Waghmare, S. Machine vision based fruit classification and grading—a review. In *ICCCE 2021* (eds Kumar, A. & Mozar, S.) (Springer Nature, 2022).
25. Miranda, J. C. et al. Fruit sizing using AI: A review of methods and challenges. *Postharvest Biol. Technol.* **206**, 112587 (2023).
26. Neupane, C. et al. Fruit sizing in orchard: A review from caliper to machine vision with deep learning. *Sensors* **23**, 3868 (2023).
27. Panchbhai, K. G. et al. Small size CNN (CAS-CNN), and modified MobileNetV2 (CAS-MODMOBNET) to identify cashew nut and fruit diseases. *Multimed. Tools Appl.* **83**, 89871–89891 (2024).
28. Panchbhai, K. G. & Lanjewar, M. G. Enhancement of tea leaf diseases identification using modified SOTA models. *Neural Comput. Appl.* **37**, 2435–2453 (2024).
29. Lanjewar, M. G. & Morajkar, P. Modified transfer learning frameworks to identify potato leaf diseases. *Multimed. Tools Appl.* **83**, 50401–50423 (2024).
30. Panchbhai, K. G., Lanjewar, M. G. & Naik, A. V. Modified MobileNet with leaky ReLU and LSTM with balancing technique to classify the soil types. *Earth Sci. Inform.* **18**, 77 (2025).
31. Yu, X., Lu, H. & Wu, D. Development of deep learning method for predicting firmness and soluble solid content of postharvest Korla fragrant pear using Vis/NIR hyperspectral reflectance imaging. *Postharvest Biol. Technol.* **141**, 39–49 (2018).
32. Liu, X., Wu, X. & Li, G. Optimized prediction of sugar content in ‘Snow’ pear using near-infrared diffuse reflectance spectroscopy combined with chemometrics. *Spectrosc. Lett.* **52**, 376–388 (2019).
33. Munawar, A. A., Zulfahrizal, Meilina & H. & Pawelzik, E.,... Near infrared spectroscopy as a fast and non-destructive technique for total acidity prediction of intact mango: Comparison among regression approaches. *Comput. Electron. Agric.* **193**, 106657 (2022).
34. Nguyen, C. N. et al. Precise sweetness grading of mangoes (*Mangifera indica* L.) based on random forest technique with low-cost multispectral sensors. *IEEE Access* **8**, 212371–212382 (2020).
35. Lanjewar, M. G., Asolkar, S. & Parab, J. S. Hybrid methods for detection of starch in adulterated turmeric from colour images. *Multimed. Tools Appl.* **83**, 65789–65814 (2024).
36. Cetinic, E., Lipic, T. & Grgic, S. Fine-tuning Convolutional Neural Networks for fine art classification. *Expert Syst. Appl.* **114**, 107–118 (2018).
37. Huang, Z., Pan, Z. & Lei, B. Transfer learning with deep convolutional neural network for SAR target classification with limited labeled data. *Remote Sens.* **9**, 907 (2017).
38. Szegedy, C. et al. Going deeper with convolutions. In *2015 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)* (ed. Szegedy, C.) 1–9 (IEEE, 2015). <https://doi.org/10.1109/CVPR.2015.7298594>.
39. Ferentinos, K. P. Deep learning models for plant disease detection and diagnosis. *Comput. Electron. Agric.* **145**, 311–318 (2018).
40. Li, W. et al. Classification and detection of insects from field images using deep learning for smart pest management: A systematic review. *Ecol. Inform.* **66**, 101460 (2021).
41. Zhou, J. et al. Graph neural networks: A review of methods and applications. *AI Open* **1**, 57–81 (2020).
42. Hu, J., Shen, L. & Sun, G. Squeeze-and-excitation networks. In *2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition* (eds Hu, J. et al.) 7132–7141 (IEEE, 2018). <https://doi.org/10.1109/CVPR.2018.00745>.
43. Woo, S., Park, J., Lee, J.-Y. & Kweon, I. S. CBAM: Convolutional block attention module. In *Computer Vision - ECCV 2018* (eds Ferrari, V. et al.) (Springer International Publishing, 2018).
44. Wang, Q. et al. ECA-Net: Efficient Channel Attention for Deep Convolutional Neural Networks. In *2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR) 11531–11539* (ed. Wang, Q.) (IEEE, 2020). <https://doi.org/10.1109/CVPR42600.2020.01155>.
45. Damaneh, M. M., Mohanna, F. & Jafari, P. Static hand gesture recognition in sign language based on convolutional neural network with feature extraction method using ORB descriptor and Gabor filter. *Expert Syst. Appl.* **211**, 118559 (2023).
46. Prabhakaran, S., Annie Uthra, R. & Preetharoselyn, J. Feature extraction and classification of photovoltaic panels based on convolutional neural network. *Comput. Mater. Contin.* **74**, 1437–1455 (2023).
47. Simonyan, K. & Zisserman, A. Very deep convolutional networks for large-scale image recognition. 3rd Int. Conf. Learn. Represent. ICLR 2015 (2015).

Acknowledgements

This research work was supported by the Taizhou Science and Technology Bureau Planning Project, China (grant No. 22NYA0).

Author contributions

Conceptualization, D.W.; methodology, C.F.; software, C.F.; Writing—original draft, C.F.; Funding acquisition, C.F. and D.W.; Resources, D.W.; Project administration, D.W.; Writing—review & editing, R.S. and M.Y.; Visualization, R.S., M.Y., Z.X. and S.D.; Supervision, R.S. and S.D.; Validation, M.Y. and W.Y.; Data Curation, Z.X. and W.Y.; All authors have read and viewed the manuscript.

Declarations

Competing interests

The authors declare no competing interests.

Additional information

Supplementary Information The online version contains supplementary material available at <https://doi.org/10.1038/s41598-025-96297-9>.

Correspondence and requests for materials should be addressed to D.W.

Reprints and permissions information is available at www.nature.com/reprints.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Open Access This article is licensed under a Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License, which permits any non-commercial use, sharing, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if you modified the licensed material. You do not have permission under this licence to share adapted material derived from this article or parts of it. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by-nc-nd/4.0/>.

© The Author(s) 2025