



OPEN ACCESS

EDITED BY

Xi Tian,
Beijing Academy of Agriculture and
Forestry Sciences, China

REVIEWED BY

Yaoyao Fan,
Shenyang Agricultural University, China
Guoxiong Zhou,
Central South University Forestry and
Technology, China
Yu Xia,
Shaanxi University of Science and
Technology, China

*CORRESPONDENCE

Xiaofei Fan
✉ leopardxf@163.com

†These authors have contributed equally to
this work

RECEIVED 29 June 2023

ACCEPTED 04 August 2023

PUBLISHED 24 August 2023

CITATION

Song C, Peng B, Wang H, Zhou Y, Sun L,
Suo X and Fan X (2023) Maize seed
appearance quality assessment based
on improved Inception-ResNet.
Front. Plant Sci. 14:1249989.
doi: 10.3389/fpls.2023.1249989

COPYRIGHT

© 2023 Song, Peng, Wang, Zhou, Sun, Suo
and Fan. This is an open-access article
distributed under the terms of the [Creative
Commons Attribution License \(CC BY\)](#). The
use, distribution or reproduction in other
forums is permitted, provided the original
author(s) and the copyright owner(s) are
credited and that the original publication in
this journal is cited, in accordance with
accepted academic practice. No use,
distribution or reproduction is permitted
which does not comply with these terms.

Maize seed appearance quality assessment based on improved Inception-ResNet

Chang Song[†], Bo Peng[†], Huanyue Wang, Yuhong Zhou,
Lei Sun, Xuesong Suo and Xiaofei Fan*

College of Mechanical and Electrical Engineering, Hebei Agricultural University, Baoding, China

Current inspections of seed appearance quality are mainly performed manually, which is time-consuming, tedious, and subjective, and creates difficulties in meeting the needs of practical applications. For rapid and accurate identification of seeds based on appearance quality, this study proposed a seed-quality evaluation method that used an improved Inception-ResNet network with corn seeds of different qualities. First, images of multiple corn seeds were segmented to build a single seed image database. Second, the standard convolution of the Inception-ResNet module was replaced by a depthwise separable convolution to reduce the number of model parameters and computational complexity of the network. In addition, an attention mechanism was applied to improve the feature learning performance of the network model and extract the best image information to express the appearance quality. Finally, the feature fusion strategy was used to fuse the feature information at different levels to prevent the loss of important information. The results showed that the proposed method had decent comprehensive performance in detection of corn seed appearance quality, with an average of 96.03% for detection accuracy, 96.27% for precision, 96.03% for recall rate, 96.15% for F1 value of reconciliation, and the average detection time of an image was about 2.44 seconds. This study realized rapid nondestructive detection of seeds and provided a theoretical basis and technical support for construction of intelligent seed sorting equipment.

KEYWORDS

corn seed, quality assessment, depthwise separable convolution, attention mechanism, feature fusion

1 Introduction

Maize is an essential cereal crop that is widely grown worldwide and has an increasing production and trade volume (Ali et al., 2020). Appearance quality is an important factor affecting the price of corn seeds, and effective identification of seed quality is critical for ensuring food security and agricultural production safety. With the rapid advancements in automation, machine vision technology (Huang et al., 2019; Kim et al., 2020; Wang and Xiao, 2020; Ansari et al., 2021; Lu et al., 2022) can be used to nondestructively and quickly

obtain seed surface feature information at a low cost and high detection accuracy and efficiency, thereby providing potential new methods for seed quality identification.

Machine learning (Rahmani et al., 2016; Prajapati et al., 2018; Liang et al., 2019; Sharaff et al., 2021) is an active field of artificial intelligence research that has advantages in terms of training small data samples and wide applications in agricultural product identification and defect detection. Farooqui et al. (2019) used a gray-level co-occurrence matrix for disease feature extraction, a support vector machine classifier for plant disease identification, combined with advanced neural network to optimize the data to improve the detection accuracy and demonstrated the feasibility of this approach for plant disease diagnosis through experiments. Effective classification of seeds is an important part of selecting and breeding good seeds. To simplify the seed selection process, Ma et al. (2021) proposed a peanut seed appearance quality detection method, using peanut seed size and appearance color as the main features and a support vector machine classification model to complete the classification task. The experimental results showed that the method had an accuracy of 86% for the classification of bulk peanut seeds, which met the preliminary classification requirements of peanut seeds in actual production. Gao et al. (2016) designed a fresh corn quality detection classifier, which analyzed the texture features of fresh corn images by wavelet analysis method, used the maximum entropy function to measure the separation degree of the texture images, and combined with the weight criterion to classify the fresh corn of different varieties, sizes, and damage degrees, and the effective classification rate could reach more than 99%. Zhao et al. (2022) extracted three categories of raw coffee bean features: contour, color, and texture to detect defective raw coffee beans by the features of a single category or category combinations. The findings were applied to a grid search to determine support vector machine classification model parameters and combined with a k-fold cross-validation test to compare support vector machine model performances. The experimental results showed that the average accuracy, precision, recall, and F1 values were 84.9%, 85.8%, 82.3%, and 84.0%, respectively. This method provided a theoretical base for the automatic detection of defective raw coffee beans. With the requirements of strict and precise agricultural development, there is an urgent need to explore new research methods to achieve precise assessment of seed appearance quality and promote intelligent agricultural development.

With the rapid development of deep learning, convolutional neural networks are widely used in the fields of medicine, aviation, and agriculture because of their excellent feature learning and expression capabilities (Kamilaris and Prenafeta-Boldu, 2018; Naranjo-Torres et al., 2020; Zhang et al., 2020; Cong and Zhou, 2022; Liu et al., 2023). Compared with traditional machine learning techniques, convolutional neural networks are more generalizable, faster to train, and can obtain significant information directly from images, which eliminates the tedious steps of manually extracting image features used in traditional methods. In applications for agriculture, convolutional neural networks are often used in areas such as the classification of crop pests and diseases (Wu et al., 2019; Peng et al., 2019; Tiwari et al., 2021; Liu et al., 2022; Liu et al., 2022), agricultural product species identification (Ajit et al., 2020; Gao

et al., 2020; Chen et al., 2021; Laabassi et al., 2021; Sj et al., 2021), yield estimation (Zhang et al., 2020; Tan et al., 2019; Alexandros et al., 2023; Kavita et al., 2023), and crop quality grading (Anikó and Miklós, 2022; Liu et al., 2022; Li et al., 2022; Wang Z. et al., 2022; Peng et al., 2023), in which they greatly promote the development of agricultural intelligence. Along with the arrival of the era of big data, the amount of image information increases exponentially, resulting in an increase in the amount of computation and training difficulty in the training process. This has also led researchers to pay more attention to lightweight networks, in order to maintain the accuracy of the premise of lightweight transformation of the network, MobileNet, ShuffleNet and other lightweight networks came into being, which can be better adapted to the evolving needs of the mobile market. To effectively alleviate the large amounts of computational resources and storage costs required for real-time image processing, Yuan et al. (2022) constructed a high-performance low-resolution MobileNet model, in which the network structure was simplified by cropping and the inception structure was used to fill the Dwise layer in a depth-separable convolution to extract the richer low-resolution features. The experimental results showed that the model achieved 89.38%, 71.60%, and 87.08% accuracies with the CIFAR - 10, CIFAR - 100, and CINIC - 10 datasets, respectively, and was suitable for real-time image classification tasks in low-resolution application scenarios. Fang et al. (2022) proposed a new network structure, HCA-MFFNet, for maize leaf disease recognition in complex contexts, and in order to validate the feasibility and effectiveness of the model in complex environments, it was compared with the existing methods, and the results proved that the model had an average recognition accuracy of 97.75% and an F1 value of 97.03%, which was the best overall performance. Hou et al. (2020) proposed a damage classification algorithm for castor seeds based on a convolutional neural network. Authors used castor seeds with missing shells or cracks and intact castor seeds to construct a dataset and build a network model to classify the seeds. The experimental results showed that the average accuracy was 87.78%, with 96.67% for castor seeds without shells, 80.00% for cracked castor seeds, and 86.67% for intact castor seeds; therefore, this method provided a feasible solution for the online real-time classification of castor seeds. Wang L. et al. (2022) designed a defect detection method based on the watershed algorithm and a two-channel convolutional neural network model, which can effectively identify defective and non-defective seeds with an average accuracy of 95.63%, an average recall of 95.29% and an F1 value of 95.46%. The assay provided an effective tool for the detection of corn seed defects. Cai et al. (2023) proposed a new grape leaf disease identification model, which was proved to have an identification accuracy of 93.26%, effectively providing decision-making information for the grape leaf disease identification system in precision agriculture. The above research results simplify the network structure complexity to a certain extent and reduced the requirement of hardware devices for model training. However, this is also often accompanied by poor model recognition accuracy, making it difficult to meet the needs of practical applications. Therefore, additional methods must be developed to improve the accuracy and achieve the purpose of accurate recognition.

In this study a seed-appearance quality assessment method based on an improved Inception-ResNet was proposed using intact and damaged maize seeds as test samples. By replacing depthwise separable convolution, adding an attention mechanism, and introducing a feature fusion strategy, the Inception-ResNet network structure was improved and optimized to obtain more detailed feature information, with the aims of achieving accurate, rapid, and nondestructive detections of seed appearance quality and providing a feasible reference scheme for subsequent automatic seed quality sorting processes.

2 Materials and methods

2.1 Experimental materials and treatment

2.1.1 Image acquisition

In this experiment, 50 groups of corn seeds with good appearance and broken appearance were collected respectively, with a total of 982 seeds, including 458 good seeds and 524 broken seeds. The training set and test set were divided according to 4:1. Ten sets of maize seed image data containing both good and defective seeds were used for verification of the final model

(Figure 1). The image acquisition platform mainly consisted of four parts: a multispectral surface array camera from JAI Company in Denmark, with an image resolution of 2048×1536; a bracket to adjust the camera height; a light source on both sides of the camera; and a shelf to place the maize seed samples. The image data acquisition device is shown in Figure 2.

2.1.2 Image processing

The Python3.6 script language was used to segment the corn seed images, as shown in Figure 3A. First, the original color image was converted to grayscale, and a binary image was then obtained using the adaptive thresholding method, whereby the seed region is shown as white and the background region as black. White noise in the image was removed using a morphological open operation, and expansion was used in the foreground to distinguish the background and foreground areas of the image. The distanceTransform function was then used to obtain the center region of each seed, and the expanded image was subtracted from the central region to obtain the edge region. Finally, the watershed algorithm was used to mark the identified central region, delineate the seed boundary, and determine the range of each seed in the complete image by the location coordinates for segmentation. The segmentation effect is shown in Figure 3B.

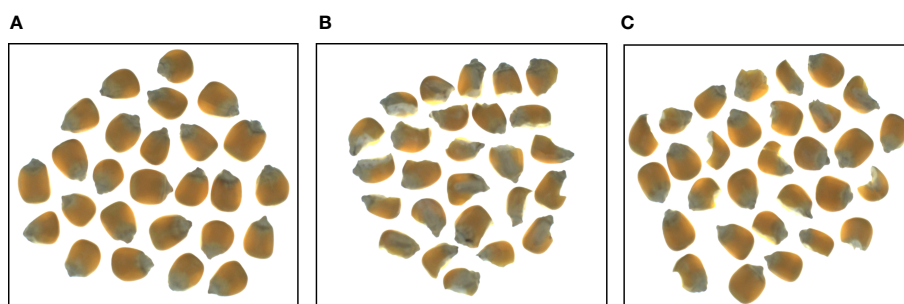


FIGURE 1
Corn seeds with different appearance qualities (A) Good seed grain, (B) Defective grain, (C) Both conditions.

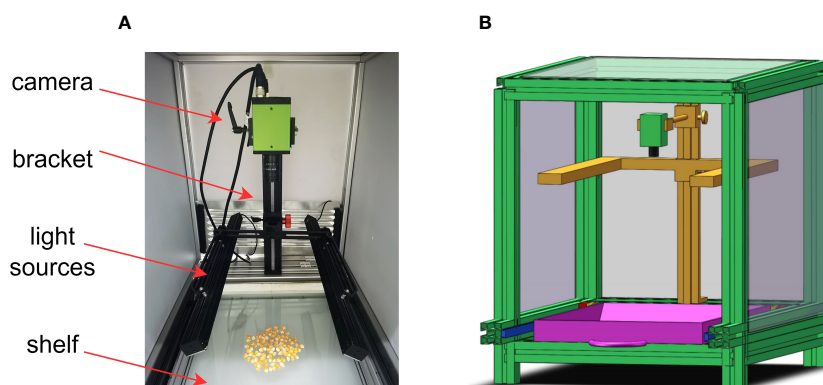


FIGURE 2
Image acquisition platform (A) physical device diagram and (B) 3D view of the device.

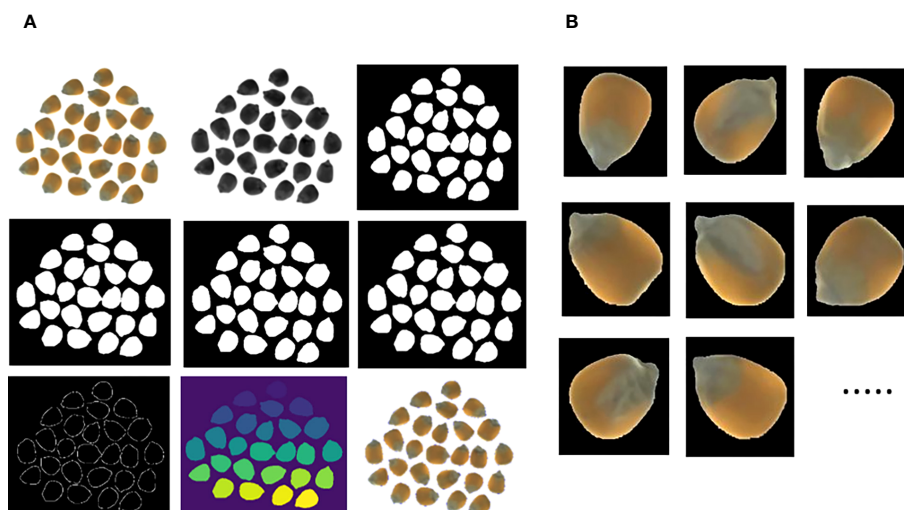


FIGURE 3 Process of watershed algorithm (A) Segmentation process and (B) Segmentation results.

2.2 Basic method and test environment

2.2.1 Convolutional neural networks

A convolutional neural network (Rong et al., 2019) is a kind of multilayer perceptron. A traditional convolutional neural network consists of an input layer, convolutional layer, pooling layer, and fully connected layer. A simple neural network model can be formed by mixing different depths and stacking orders, as shown in Figure 4, in which the term random represents the number of times a particular layer of a structure is randomized.

The input layer represents the initial input of the entire convolutional neural network, and in the case of image processing, the input layer contains the pixel matrix of the input image, and the

convolutional layer is the core component of convolutional neural network, which has the characteristics of local connection, weight sharing and translation invariance, and realizes the feature extraction function of the network. The pooling layer compresses the size of the input feature map while extracting the main features, thereby reducing the complexity of the network training process and the probability of overfitting the network model during this process. The fully connected layer is involved in weighting the features extracted from the convolutional and pooling layers, transforming them into layer vectors, and inputting one-dimensional data into the Softmax layer through multiplication operations to obtain the image classification results. It can be understood as a parallel, large-scale distributed processor that is capable of storing and using empirical knowledge.

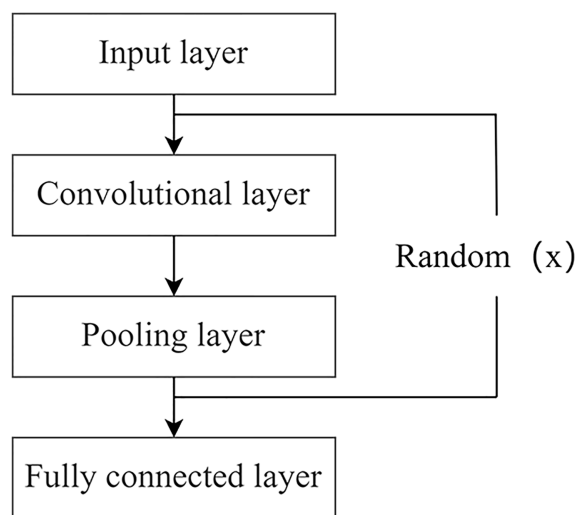


FIGURE 4 Basic structure of convolutional neural network.

2.2.2 Depthwise separable convolution

Depthwise separable convolution (Mamalet and Garcia, 2012) was proposed as a typical lightweight convolution structure that has a significantly reduced parameter number and increased training speed compared to a standard convolution and can separate channels and regions in the convolution operation. As shown in Figure 5, depthwise separable convolution mainly consists of two parts, depthwise convolution and pointwise convolution, which are used to extract feature information. One depthwise convolution kernel is responsible for one channel, and the number of feature maps generated by this process equals the number of input channels. The convolution operation is performed independently for each channel of the input layer, which does not make full use of the feature information from other channels in the same spatial location. Therefore, pointwise convolution is required to combine the feature maps to generate a new feature map, in which the number of convolution kernels corresponds to the output feature map.

Depthwise separable convolution effectively reduces the number of parameters required for the network model compared to the normal convolutional approach. The $N \times H \times W \times C$ convolution kernels can be replaced by $C \times H \times W \times 1$ depthwise and $N \times 1 \times 1 \times C$ pointwise convolution kernels. The number of depthwise convolution parameters is $(H \times W \times 1) \times C$, and the number of pointwise convolution parameters is $(1 \times 1 \times C) \times N$. The combined number of parameters for the depthwise separable convolution can be calculated as follows:

$$Params = H \times W \times C + C \times N \tag{1}$$

The number of parameters in the ordinary convolution is $H \times W \times C \times N$, and the relationship between the two is compared as follows:

$$\frac{H \times W \times C + C \times N}{H \times W \times C \times N} = \frac{1}{N} + \frac{1}{H \times W} \tag{2}$$

2.2.3 Attention mechanism

The working principle of the attention mechanism is similar to the selective attention of human vision. It utilizes limited computational resources to focus on important feature information and emphasize regions of interest in a dynamically

weighted manner, discarding irrelevant background information and nuisance information in the input features to improve network performance. In general, the attention mechanism involves the process of weight assignment, in which the input feature information is processed, the attention information is obtained through a weight assignment, and the attention mechanism is used to process these features. The process can be expressed as

$$Attention = f(g(x), x) \tag{3}$$

Here, x represents the input feature information of the attention mechanism; $g(x)$ represents the focus on the key areas, that is, the process of generating attention information by processing the input features; and $f(g(x), x)$ represents the processing of the key areas, that is, processing of the input information using the attention information obtained from $g(x)$.

2.2.4 Evaluation indicators

In this study, objective evaluation criteria (Feng et al., 2022; Wang X. et al., 2022) were used to analyze the inspection model of the maize seed appearance quality using Accuracy (A), Precision (P), Recall (R), and by introducing the F1 value as the average evaluation of the reconciliation. The related formulae are as follows:

$$A = \frac{TP + TN}{TP + TN + FN + FP} \times 100\% \tag{4}$$

$$P = \frac{TP}{TP + FP} \times 100\% \tag{5}$$

$$R = \frac{TP}{TP + FN} \times 100\% \tag{6}$$

$$F1 = \frac{2PR}{P + R} \times 100\% \tag{7}$$

2.2.5 Test environment

We used a Windows 10, 64-bit operating system with a x64 based processor, Cuda version 11.0, and the Tensorflow deep learning framework based on the Python programming language. The computer contained an NVIDIA GeForce RTX 3090 graphics card with 24G video memory and a 12th Gen Intel(R) Core (TM) i7-12700KF processor at 3.61 GHz.

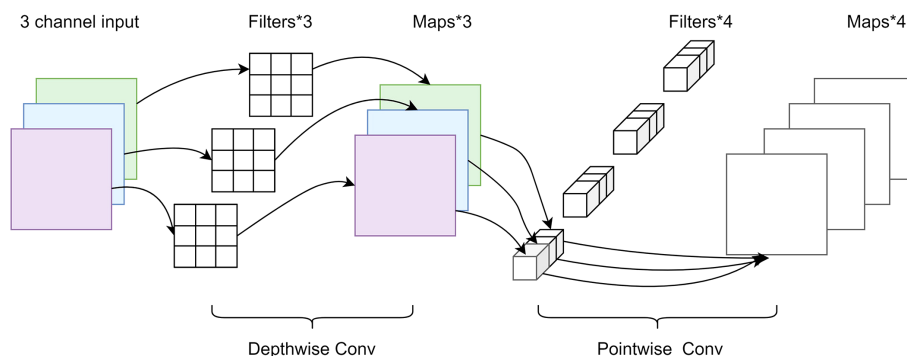


FIGURE 5
Depthwise separable convolution.

2.3 Algorithm improvement

2.3.1 Improving the Inception-ResNet module

The traditional Inception module, shown in Figure 6A, consists of 1×1 , 3×3 , and 5×5 convolution operations of various sizes for feature extraction using multiple-scale parallel convolution operations. The Inception-ResNet module (Figure 6B) combined the residual network structure of the ResNet with the separation of the large convolutional network into two tandem small convolutional structures to obtain the output feature maps of the 5×5 convolution, thereby improving the classification performance of the model. In this study, based on the Inception-ResNet module, the standard convolution part of the Inception-ResNet structure was replaced by a cost-effective convolution operation by combining the lightweight structure depthwise separable convolution (Dinception-ResNet). This approach could reduce the number of model parameters, increase the depth of the network, and enhance the feature extraction capability of the model, while preserving the feature diversity of the traditional inception multigroup structure.

A comparison of the parameters of the improved model is presented in Table 1. The original Inception-ResNet model used up to 28 979 618 parameters; this number was reduced by 177 920 after the depthwise separable convolutional replacement and the ratio of the number of trainable parameters to the total number of parameters was improved, which favors the design of lightweight networks.

2.3.2 Adding attention mechanism

Introducing the Efficient Channel Attention Network after the Dinception-ResNet module effectively avoided dimensionality reduction and captured cross-channel interactions, as shown in Figure 7. The Efficient Channel Attention Network first transformed the input feature map from a matrix $[H, W, C]$ into a vector $[1, 1, C]$ using global averaging pooling, after which it calculated the adaptive 1D convolution kernel size based on the number of channels of the feature map, which was then used in 1D convolution. Subsequently, the weights of the feature maps with respect to each channel were obtained. Finally, the normalized weights and the original input feature maps were multiplied channel-by-channel to generate the weighted feature maps. These maps can be used to solve the information overload problem and improve the efficiency and accuracy of task processing by focusing on the more critical information for the current task, reducing the attention to other information, and using a small number of parameters to achieve suitable results. To fully demonstrate the effectiveness of Efficient Channel Attention Network, we conducted a comparison experiment between it and the commonly used attention mechanisms CBAM, SENet and CANet. By testing the relevant indexes of the experiments, the results were shown in Table 2, which clearly concluded that the overall detection performance of Efficient Channel Attention Network was optimal.

2.3.3 Introducing the feature fusion mechanism

As the depth of the network model increases, important feature information may be missed in the feature extraction process,

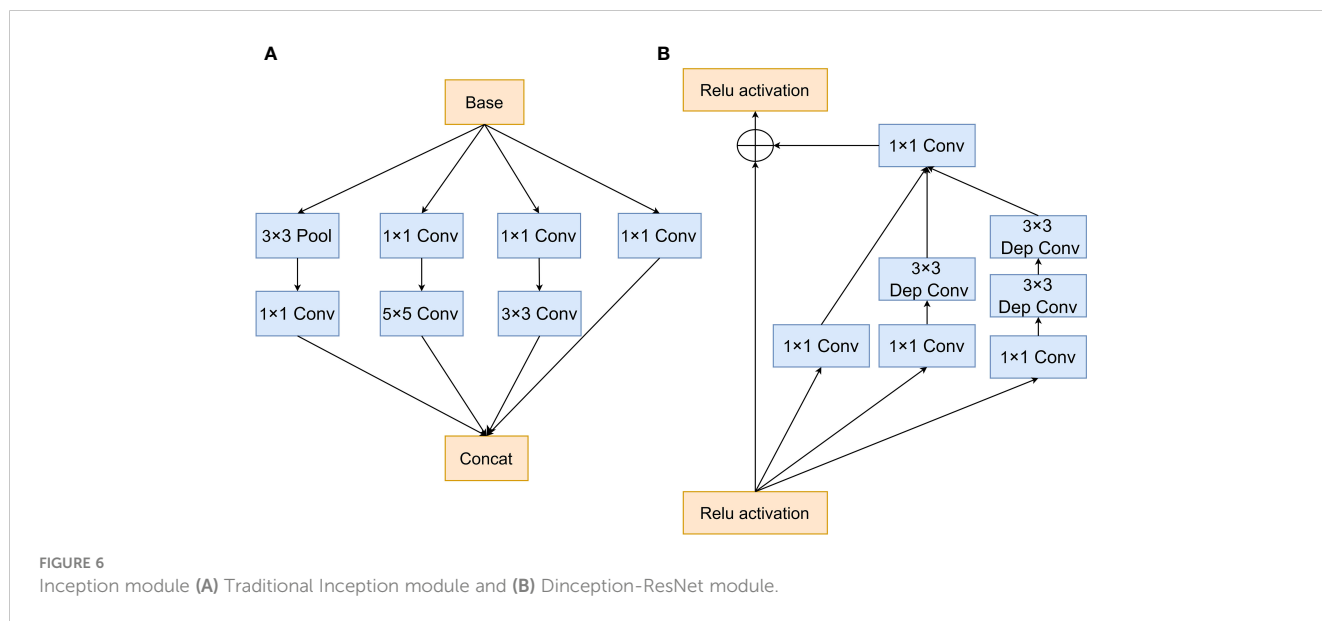


FIGURE 6 Inception module (A) Traditional Inception module and (B) Dinception-ResNet module.

TABLE 1 Parameter comparison.

Model	Total parameters	Trainable parameters	Non-trainable parameters
Inception-ResNet	28 979 618	28 947 106	32 512
Dinception-ResNet	28 801 698	28 770 338	31 360

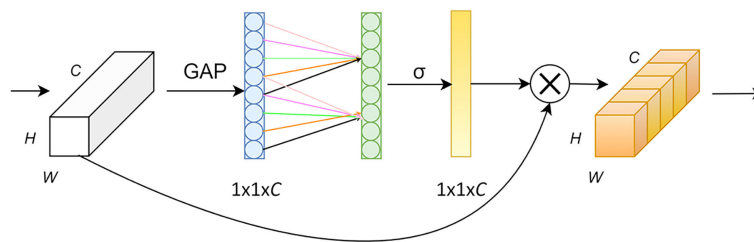


FIGURE 7
ECA network structure.

thereby decreasing the classification accuracy of the model. Therefore, this study proposed feature fusion, as shown in Figure 8, in which the output features of each module were first pooled by global averaging to avoid overfitting. Subsequently, the output feature maps of the three stages were feature-fused to enrich the output feature information, enhance the generalization performance of the model, and improve the expression ability of the feature information.

2.3.4 Improved network model structure

The traditional convolutional neural network structure generates a large number of parameters in the deep network. In addition, key features are mixed with irrelevant features, and the network is difficult to optimize; therefore, this study proposed an improved model based on Inception-ResNet (Figure 9). Depthwise separable convolution was used to replace the standard convolution in the Inception-ResNet module to reduce the number of model parameters, the network model was optimized by introducing the Efficient Channel Attention mechanism to increase the feature weights of key information and improve the network performance; in addition, the output features of the low, middle and high layers of the model were fused to improve the feature extraction ability of the model and enrich feature information to achieve the purpose of network optimization.

2.3.5 Ablation study

In order to validate the effectiveness of the method proposed in this paper, separate experiments were conducted for the proposed modules to compare with the original Inception-Resnet model. By comparing the results of groups 1-2-3-4 in Table 3, it can be seen that the performance of each module reference is improved to some extent compared to the original network. Among them, the

introduction of the feature fusion module resulted in the most obvious model performance improvement, with each performance index improved by 3.38 – 5.55 percentage points respectively. The average time for detecting an image was 2.62 seconds. Replacing standard convolution with depthwise separable convolution, the average time to detect an image was reduced by 30 milliseconds. By comparing the results of 1-2-5-6 groups and referring modules one by one, the model performance was improved and the effectiveness of the modules is fully proved.

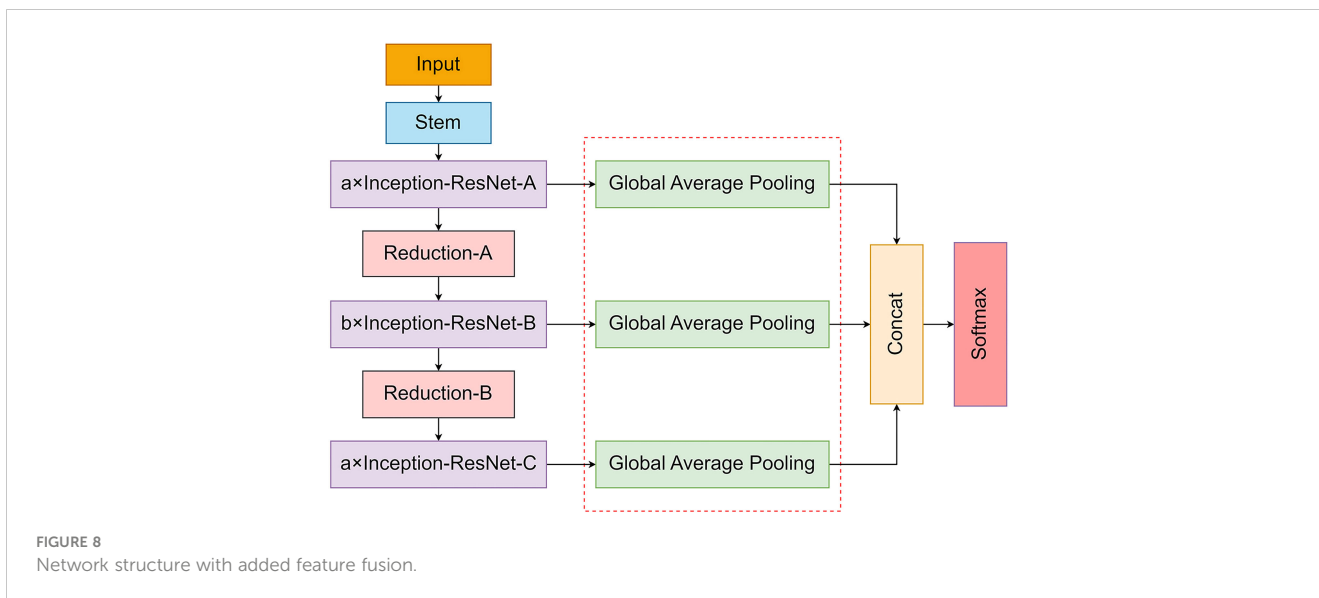
3 Results and analysis

3.1 Model optimization

Different networks apply different model parameters. By adjusting certain parameters in the model and using test set recognition accuracy as the evaluation index, we investigated the effects of the parameters on the classification accuracy of the model. In the convolutional neural network model, batch size is an important hyperparameter, and we chose a batch size that was appropriate to train the model to converge to the global optimum. Using a large number of parameters in a deep convolutional neural network creates correction issues; therefore, a suitable optimizer was selected to improve the model training speed and accuracy. The training parameter batch sizes of the model were 16, 32, and 64, and two different optimization algorithms, Adam and SGD, were used. The number of model iterations was set to 100 and the learning rate was 0.001; the final test set accuracy variation curve is shown in Figure 10. These settings ensured that when the Adam optimizer was selected for the improved model and the bath size was set to 32, the accuracy of the test set was the highest and the detection performance was optimal.

TABLE 2 Comparison of detection performance of different attention mechanisms.

Attention mechanism	Average Accuracy/%	Average Precision/%	Average Recall/%	F1/%	Detection time/s
ECANet	95.96	92.46	91.26	91.86	2.68
CBAM	89.68	91.33	89.68	90.50	2.65
SENet	83.33	87.34	83.33	85.29	2.67
CANet	90.48	91.89	90.47	91.17	2.66



3.2 Comparison of machine learning experiments

To verify the performance of the improved model, three features of maize seed (shape, color, and texture) were extracted using traditional machine vision techniques, and the three types were combined to quantify the images more effectively. The logistic regression, k-nearest neighbor, decision tree, random forest, gaussian naïve bayes, and support vector machine models were used for classification, and the common lightweight convolutional neural networks DenseNet, MobileNet, ShuffleNet, and Inception, and the original Inception-ResNet were selected for comparison. The dataset was divided into training and test sets in a 4:1 ratio, and the classification accuracies of the different models on the test set are shown in Figure 11A. The highest accuracy with the test set occurred

with the improved Inception-ResNet model, with 99.49%, which was an increase of 2.03 percentage points over the accuracy of the original model of 97.46% and an increase of 5.07 percentage points compared to the LR model, which had the highest machine-learning recognition accuracy for the test set of 94.42%. In addition, a confusion matrix was chosen as the visual presentation tool to evaluate the quality of the classification models, and the matrices for several models of the test set are shown in Figure 11B. The plots showed that the original Inception-ResNet model had an accuracy of 96.67% for the identification of seeds with a defective appearance and 98.13% for the identification of seeds with a good appearance. The confusion matrix for the improved Inception-Resnet model showed the best results, with a significant improvement and a classification accuracy of the test set seeds of up to 100%; therefore, this model provided an accurate identification of seed appearance quality.

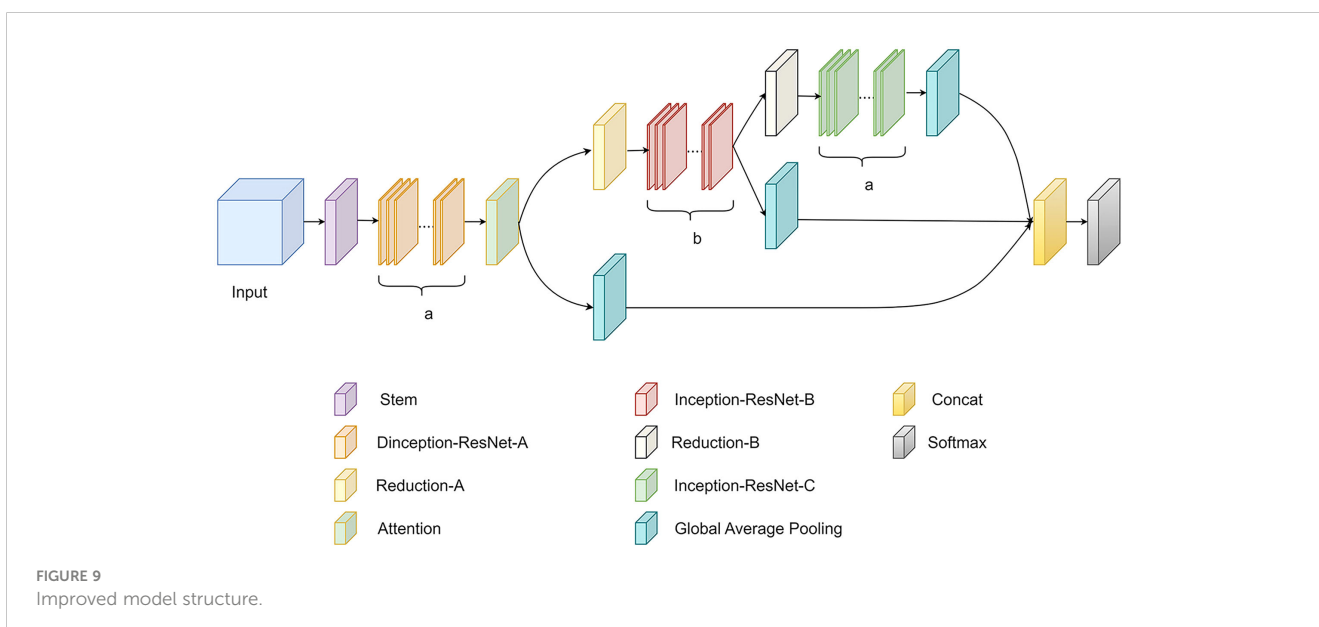


TABLE 3 Comparison of improving module performance.

Model	Average Accuracy/%	Average Precision/%	Average Recall/%	F1/%	Detection time/s
1 Inception-Resnet	88.10	90.26	88.10	89.17	2.59
2 Inception-Resnet (+depthwise separable convolution)	89.68	91.33	89.67	90.49	2.56
3 Inception-Resnet (+ECANet)	95.96	92.46	91.26	91.86	2.68
4 Inception-Resnet (+feature fusion)	93.65	94.28	93.65	93.96	2.62
5 Inception-Resnet (+depthwise separable convolution + ECANet)	92.06	93.05	92.06	92.55	2.77
6 Improved Inception-ResNet	96.03	96.27	96.03	96.15	2.44

3.3 Comprehensive testing performance

To evaluate the effectiveness of the improved Inception-ResNet in detecting seed appearance quality, ten sets of image data with good and bad co-existing seeds were selected for the final validation and performance of the six detection algorithms, DenseNet, MobileNet, ShuffleNet, Inception, Inception-ResNet, and improved Inception-ResNet. The models were evaluated using the criteria of average accuracy, precision, recall and the reconciled average evaluation F1 value. The results are shown in Table 4, and the highest average accuracy of Inception-ResNet after improvement reached 96.03%, which was 3.23 – 11.11 percentage points higher than those of the other models, the average precision reached 96.27%, an improvement of 3.46 – 8.00 percentage points compared to other networks. Similarly, the average recall was 3.22 – 11.11 percentage points greater than the other models, with 96.03%, while the reconciled average evaluation F1 value reached 96.15%, which was 3.34 – 9.59 percentage points higher than the other algorithms and the average time to achieve real-time detection was 2.44 seconds per detected image.

3.4 Model performance

Combined with the watershed algorithm to obtain the specific location coordinates of the corn seeds for defect detection, the detection results of several models are shown in Figure 12. The improved Inception-ResNet model had a significantly greater number of cases with target confidence levels of 1.0, with approximately 88% of the seed identifications showing a level approaching 1.0, which is a six-percentage point improvement over the level of the original model. A comparison of the detection results is shown in Table 5, which shows that the improved Inception-ResNet model had the greatest number of correct detections and highest overall recognition accuracy, thereby allowing for an accurate detection of the appearance quality of corn seeds to achieve the desired results. In the actual detection process, there are also cases of misdetection, such as the possibility of detecting bad seed as good seed when the area of damage is small. When the seeds are randomly placed at an inappropriate Angle or there is serious adhesion leading to area coverage between seeds, it is possible to detect good seeds as bad seeds. This also indicates that there is room for model refinement and we need to follow up to explore more accurate detection models.

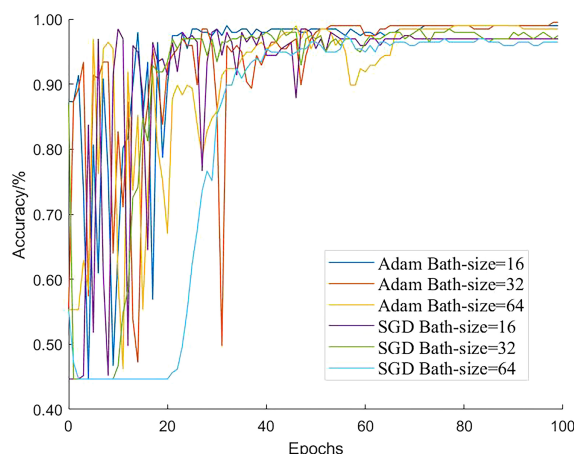


FIGURE 10 Comparison of parameters.

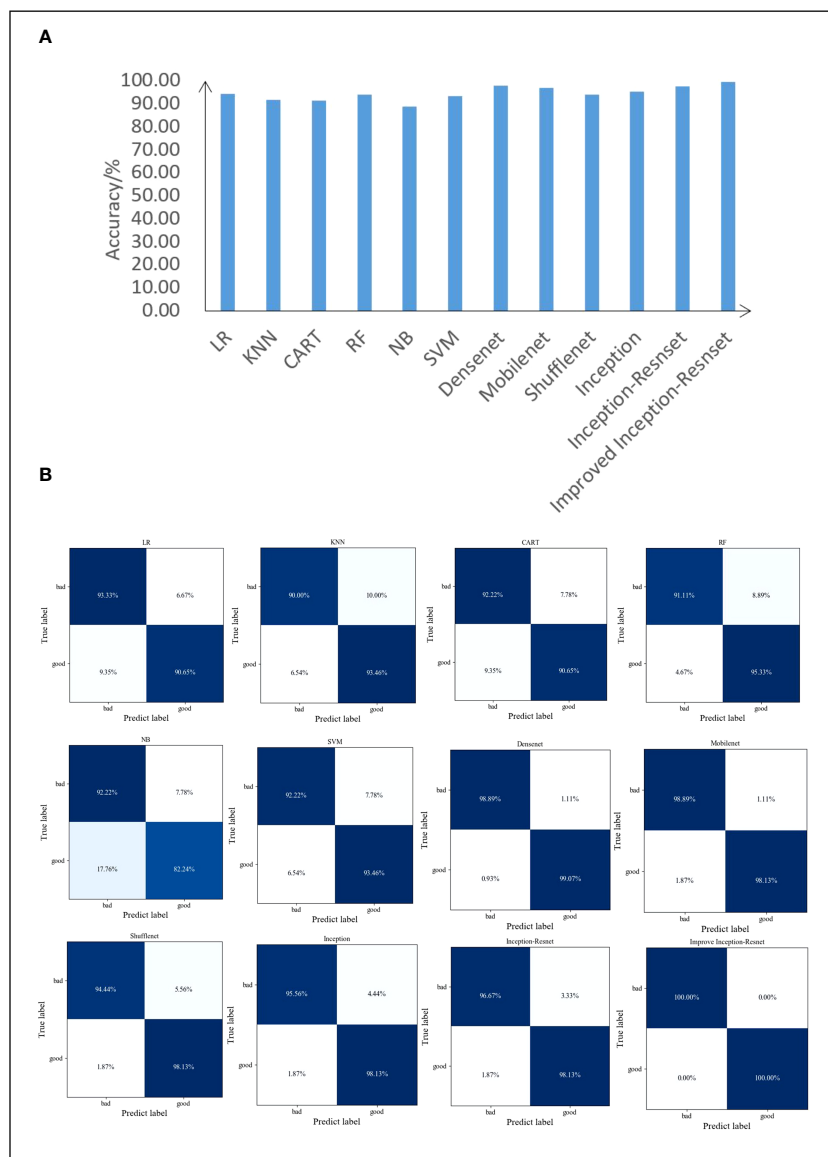


FIGURE 11 Test result (A) Comparison of accuracy between different model test sets and (B) Confusion matrix of different models.

4 Discussion

We proposed an improved model based on Inception-Resnet for corn seed appearance quality detection, which can realize the purpose of accurate and real-time detection. However, the model is domain dependent and mainly focuses on the detection of seed appearance quality, which makes it difficult to directly migrate the network to other domains. This is mainly due to the fact that the network structure modifications are performed based on seed features, which are difficult to guarantee that they will respond well on datasets from other domains. With the development of agricultural digitization, it can be applied to seed intelligent sorting equipment to guarantee the quality of seeds. We can do further research to solve some problems, firstly, the data collection was conducted indoors, in the future, we can try to train the model

directly for the dataset collected in the actual outdoor environment to improve the generalization ability of the model; secondly, we can design a lighter network structure, which can ensure the stability of the algorithm and efficiently deploy it on different platforms while pursuing high performance; finally, we only study the appearance damage of corn seeds, and in the future we can continue to explore more specific defects on the seed surface, such as diseases and pests, and further promote the development of the seed industry.

5 Conclusions

To improve the accuracy of seed appearance quality assessment, this study proposed an improved Inception-ResNet model based on the Inception-ResNet algorithm for identifying appearance defects

TABLE 4 Comparison of detection model performances.

Model	Average Accuracy/%	Average Precision/%	Average Recall/%	F1/%	Detection time/s
DenseNet	84.92	88.27	84.92	86.56	2.86
MobileNet	87.30	89.75	87.30	88.51	1.81
ShuffleNet	88.87	90.24	88.88	89.55	1.91
Inception	92.80	92.81	92.81	92.81	2.55
Inception-ResNet	88.10	90.26	88.10	89.17	2.59
Improved Inception-ResNet	96.03	96.27	96.03	96.15	2.44

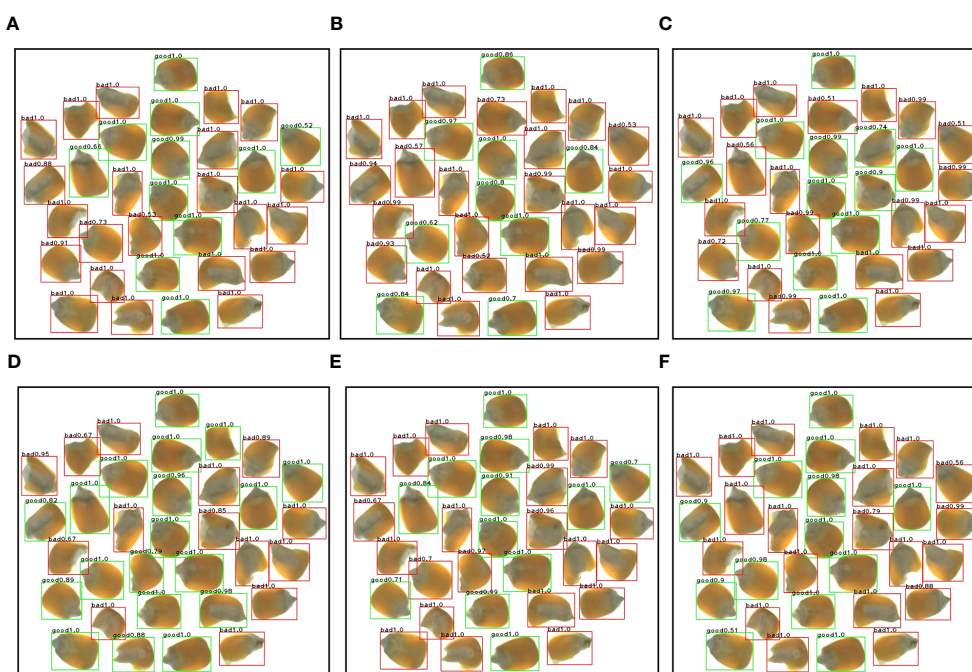


FIGURE 12 Comparison of detection results of multiple algorithms (A) DenseNet (B) MobileNet (C) ShuffleNet (D) Inception (E) Inception-ResNet (F) Improved Inception-ResNet.

of corn seeds and obtained the following conclusions: (1) Taking advantage of the small number of depthwise separable convolutional parameters as opposed to the standard convolution

in the Inception-ResNet module reduced the large number of parameters generated by overlaying the Inception-ResNet module and the requirement for hardware resources. (2) Introducing the Efficient Channel Attention Network strengthened the ability to learn key information and avoided the problem of excessive information storage and information overload in the model. Simultaneously, the output feature maps were fused to obtain richer feature information to enhance the network generalization ability and improve network performance. (3) The detection effect of the method proposed in this study was superior to the other models tested, with an average accuracy of 96.03%, average precision of 96.27%, average recall of 96.03%, F1 value of 96.15%, and detection speed for a single corn seed image of approximately 2.44 seconds. The performance index improved significantly with high performance stability, providing a theoretical basis for subsequent seed quality detection.

TABLE 5 Comparison of test results.

Model	Correct number	Miss number
DenseNet	27	6
MobileNet	25	8
ShuffleNet	25	8
Inception	27	6
Inception-ResNet	28	5
Improved Inception-ResNet	29	4

Data availability statement

The raw data supporting the conclusions of this article will be made available by the authors, without undue reservation.

Author contributions

CS, BP, and XF conceived the idea, proposed the method. HW and YZ set up the platform and collected data. CS and BP write the code, test the results and write the manuscript. LS and XS contributed to the validation results. XF and XS provided guidance for the writing of this paper. All authors contributed to the article and approved the submitted version.

Funding

This study was supported by the National Natural Science Foundation of China (32072572), the earmarked fund for CARS (CARS-23), the Talent Recruiting Program of Hebei Agricultural University (YJ2021056), and the research project on Basic Research Business Expenses of Provincial Higher Education Institutions in Hebei Province (KY2022019).

References

- Ajit, A., Acharya, K., and Samanta, A. (2020). A review of convolutional neural networks[C]//2020 international conference on emerging trends in information technology and engineering (ic-ETITE) (Vellore, India: IEEE 2020 International Conference on Emerging Trends in Information Technology and Engineering (ic-ETITE)). doi: 10.1109/ic-ETITE47903.2020.049
- Alexandros, O., Cagatay, C., and Ayalew, K. (2023). Deep learning for crop yield prediction: a systematic literature review. *New Z. J. Crop Hortic. Sci.* 51 (1), 1–26. doi: 10.1080/01140671.2022.2032213
- Ali, A., Qadri, S., Mashwani, W. K., Brahim, B., Naeem, S., Rafique, S., et al. (2020). Machine learning approach for the classification of corn seed using hybrid features. *Int. J. Food Properties* 23 (1), 1097–1111. doi: 10.7779/ijfsnt.2020.40.1.47
- Anikó, N., and Miklós, N. (2022). Crop yield prediction in precision agriculture. *Agronomy* 12 (10), 2460–2460. doi: 10.3390/agronomy13071812
- Ansari, N., Ratri, S. S., Jahan, A., Ashik-E-Rabbani, M., and Rahman, A. (2021). Inspection of paddy seed varietal purity using machine vision and multivariate analysis. *J. Agric. Food Res.* 3, 100109. doi: 10.1016/j.jafr.2021.100109
- Cai, C., Wang, Q., Cai, W., Yang, Y. X., Hu, Y. H., Li, L. J., et al. (2023). Identification of grape leaf diseases based on VN-BWT and Siamese DWOAM-DRNet. *Eng. Appl. Artif. Intell.* 123 (PB). doi: 10.1016/j.engappal.2023.106341
- Cong, S., and Zhou, Y. (2022). A review of convolutional neural network architectures and their optimizations. *Artif. Intell. Rev.* 56 (3), 1905–1969. doi: 10.3390/app10103443
- Chen, Y. H., An, X. S., Gao, S. M., Li, S. J., and Kang, H. W. (2021). A deep learning-based vision system combining detection and tracking for fast on-line citrus sorting. *Front. Plant Sci.* 12. doi: 10.3389/fpls.2021.622062
- Fang, S., Wang, Y., Zhou, G., Chen, A., Cai, W., Wang, Q., et al. (2022). Multi-channel feature fusion networks with hard coordinate attention mechanism for maize disease identification under complex backgrounds. *Comput. Electron. Agric.* 203, 107486. doi: 10.1016/j.compag.2022.107486
- Farooqui, N. A., and Ritika, R. (2019). An identification and detection process for leaves disease of wheat using advance machine learning techniques. *Bioscience Biotechnol. Res. Commun.* 12 (4), 1081–1091. doi: 10.21786/bbrc/12.4/31
- Feng, X., Zhao, C., Wang, C., Wu, H., Miao, Y., and Zhang, J. (2000). A vegetable leaf disease identification model based on image-text cross-modal feature fusion *frontiers in plant science*, (Lausanne, Switzerland: Frontiers in Plant Science) 13, 918940. doi: 10.3389/fpls.2022.918940
- Gao, X. H., and Liu, B. (2016). Design and experiment of fresh corn quality detection classifier based on machine vision. *Trans. Chin. Soc. Agric. Eng.* 32 (01), 298–306. doi: 10.11975/j.issn.1002-6819.2016.01.041
- Gao, F. F., Fu, L. S., Zhang, X., Majeed, Y., Li, R., Karkee, M., et al. (2020). Multi-class fruit-on-plant detection for apple in SNAP system using Faster R-CNN. *Comput. Electron. Agric.* 176. doi: 10.3389/fpls.2021.622062
- Hou, J. M., Yao, E. C., and Zhu, H. J. (2020). Classification of castor seed damage based on convolutional neural network. *Trans. Chin. Soc. Agric.* 51 (S1), 440–449. doi: 10.6041/j.issn.1000-1298.2020.S1.052
- Huang, S., Fan, X. F., Sun, L., Shen, Y. L., and Suo, X. S. (2019). Research on classification method of maize seed defect based on machine vision. *J. Sensors* 2019, 1–9. doi: 10.1155/2019/2716975
- Kamilaris, A., and Prenafeta-Boldu, F. X. (2018). A review of the use of convolutional neural networks in agriculture. *J. Agric. Sci.* 156 (3), 312–322. doi: 10.1016/j.compag.2022.107605
- Kavita, J., Prathista, M., Sanchit, J., and Sukriti, N. (2023). Crop yield prediction using machine learning and deep learning techniques. *Proc. Comput. Sci.* 218, 406–417. doi: 10.1016/j.procs.2023.01.023
- Kim, D. H., Boo, S. B., Hong, H. C., Yeo, W. G., and Lee, N. Y. (2020). Machine vision-based defect detection using deep learning algorithm. *J. Korean Soc. Nondestructive Testing* 40 (1), 47–52. doi: 10.7779/ijksnt.2020.40.1.47
- Laabassi, K., Belarbi, M. A., Mahmoudi, S., Ahmed, M. S., Ferhat, K., et al. (2021). Wheat varieties identification based on a Deep Learning approach. *J. Saudi Soc. Agric. Sci* 20 (5), 281–289. doi: 10.1016/j.jssas.2021.02.008
- Li, L., Hu, D. Y., Tang, T. Y., and Tang, Y. L. (2022). Nondestructive testing of the quality of different grades of creamy strawberries based on Laida algorithm. *J. Food Process. Preservation* 46 (11). doi: 10.1111/JFPP.17008
- Liang, W. J., Zhang, H., Zhang, G. F., and Gao, H. X. (2019). Rice blast disease recognition using a deep convolutional neural network. *Sci. Rep.* 9 (1), 1–10. doi: 10.1038/s41598-019-38966-0
- Liu, S., Qiao, Y. L., Li, J. W., Zhang, H. T., Zhang, M. K., and Wang, M. L. (2022). An improved lightweight network for real-time detection of apple leaf diseases in natural scenes. *Agronomy* 12 (10). doi: 10.3390/AGRONOMY12102363
- Liu, J., Wang, X., Miao, W., and Liu, G. X. (2022). Tomato pest recognition algorithm based on improved YOLOv4. *Front. Plant Sci.* 13, 814681. doi: 10.1109/ACCESS.2019.2914929

Acknowledgments

Thanks to the teachers in the project room for their guidance and the help of the students. And we would like to thank Editage (www.editage.cn) for English language editing.

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

- Liu, M. Y., Zhou, G. X., Chen, A. B., Li, L. J., and Hu, Y. H. (2023). Identification of tomato leaf diseases based on LMBRNet. *Eng. Appl. Artif. Intell.* 123 (PA).
- Lu, C. Q., Li, Z. Y., Shi, S. Y., and Peng, S. L. (2022). Research on target defect detection based on machine vision. *J. Physics: Conf. Ser.* 2219 (1). doi: 10.1088/1742-6596/2219/1/012060
- Ma, J. J., Wang, K. Q., Zheng, Y. X., Cai, K., and Lin, Q. Y. (2021). Research on the appearance quality detection and classification of peanut seeds based on machine vision. *Anhui Agric. Sci.* 49 (10), 225–227, 231. doi: 10.11975/j.issn.1002-6819.2022.14.033
- Mamalet, F., and Garcia, C. (2012). *Simplifying convnets for fast learning[C]//International Conference on Artificial Neural Networks* (Berlin, Heidelberg: Springer), 58–65.
- Naranjo-Torres, J., Mora, M., Hernández-García, R., Barrientos, R. J., Fredes, C., and Valenzuela, A. (2020). A review of convolutional neural network applied to fruit image processing. *Applied Sciences* 10 (10), 3443. doi: 10.3390/app10103443
- Peng, J., Chen, Y. H., Liu, B., He, D. J., and Liang, C. Q. (2019). Real-Time detection of apple leaf diseases using deep learning approach based on improved convolutional neural networks. *IEEE Access* 7, 59069–59080. doi: 10.1109/ACCESS.2019.2914929
- Peng, D., Li, W., Zhao, H., Zhou, G., and Cai, C. (2023). Recognition of tomato leaf diseases based on DIMPCNET. *Agronomy* 13, 1812. doi: 10.3390/agronomy13071812
- Prajapati, H., Shah, J., and Dabhi, V. (2018). Detection and classification of rice plant diseases. *Intelligent Decision Technol.* 2), 1–18. doi: 10.3233/idt-180338
- Rahmani, M. E., Amine, A., and Hamou, R. M. (2016). Supervised machine learning for plants identification based on images of their leaves. *Int. J. Agric. Environ. Inf. Syst.* 7 (4), 17–31. doi: 10.4018/IJAEIS.20211001.OA5
- Rong, D., Xie, L. J., and Ying, Y. B. (2019). Computer vision detection of foreign objects in walnuts using deep learning. *Comput. Electron. Agric.* 162, 1001 – 1010. doi: 10.1016/j.compag.2019.05.019
- Sj, A., Shma, B., Fjv, A., and Am, C. (2021). Computer-vision classification of corn seed varieties using deep convolutional neural network. *J. stored products Res.* doi: 10.1016/j.jspr.2021.101800. 92, 101800. doi: 10.1016/j.jspr.2021.101800
- Sharaff, A., and Patel, B. (2021). Rice crop disease prediction using machine learning technique. *Int. J. Agric. Environ. Inf. Syst.* 12 (4), 1–15. doi: 10.4018/IJAEIS.20211001.OA5
- Tan, S. Y., Ma, X., Mai, Z. J., Qi, L., and Wang, Y. W. (2019). Segmentation and counting algorithm for touching hybrid rice grains. *Comput. Electron. Agric.* 162, 493–504. doi: 10.1016/j.compag.2019.04.030
- Tiwari, V., Joshi, R. C., and Dutta, M. K. (2021). Dense convolutional neural networks based multiclass plant disease detection and classification using leaf images. *Ecol. Inf.* doi: 10.1016/j.ecoinf.2021.101289
- Wang, L., Liu, J., Zhang, J., Wang, J., and Fan, X. (2022). Corn seed defect detection based on watershed algorithm and two-pathway convolutional neural networks. *Front. Plant Sci.* 13. doi: 10.3389/fpls.2022.730190
- Wang, Z. L., Huang, W. Q., Tian, X., Long, Y., Li, L. J., and Fan, S. X. (2022). Rapid and Non-destructive Classification of New and Aged Maize Seeds Using Hyperspectral Image and Chemometric Methods. *Front. Plant Sci.* 13. doi: 10.3389/FPLS.2022.849495
- Wang, Z., and Xiao, Z. (2020). *Product quality detection and recognition based on vision and deep learning[C]//2020 35th youth academic annual conference of chinese association of automation (YAC)* (China: Youth Academic Annual Conference of Chinese Association of Automation (YAC)). doi: 10.1109/yac51587.2020.9337602
- Wang, X., Yang, W., Lv, Q., Huang, C., Liang, X., Chen, G., et al (2022). Field rice panicle detection and counting based on deep learning. *Front. Plant Sci.* 13, 966495. doi: 10.3389/fpls.2022.966495
- Wu, J. Y., Li, B., and Wu, Z. L. (2019). Detection of crop pests and diseases based on deep convolutional neural network and improved algorithm[P]. *Mach. Learn. Technol.*, 20–27. doi: 10.1145/3340997.3341010
- Yuan, H. Y., Cheng, J. P., Wu, Y. R., and Zeng, Z. Y. (2022). Low-res MobileNet: An efficient lightweight network for low-resolution image classification in resource-constrained scenarios. *Multimedia Tools Appl.* doi:10.1007/S11042-022-13157-8. 81 (27), 38513–38530. doi: 10.1007/S11042-022-13157-8
- Zhang, C., Zhang, Y. Y., Yan, T. Y., Bai, X. L., Xiao, Q. L., Gao, P., et al. (2020). Application of near-infrared hyperspectral imaging for variety identification of coated maize kernels with deep learning. *Infrared Phys. Technol.* 111 (20), 103550. doi: 10.1016/j.compag.2019.04.030
- Zhao, Y. Q., Yang, H. L., Zhang, Y., Yang, Y. K., Yang, Y., and Sai, M. (2022). Detection of defective Arabica green coffee beans based on feature combination and SVM. *Trans. Chin. Soc. Agric. Eng. (Transactions CSAE)* 38 (14), 295–302. doi: 10.11975/j.issn.1002-6819.2022.14.033