IEEE Access
Multidisciplinary | Rapid Review | Open Access Journal

Date of publication xxxx 00, 0000, date of current version xxxx 00, 0000.

Digital Object Identifier 10.1109/ACCESS.2022.Doi Number

Hyperspectral RGB Imaging Combined with Deep Learning for Maize Seed Variety Identification

Jian Li[1,3], Fan Xu[1,3], Shaozhong Song[2,*] , Qi Ji[4] and Junling Liu[2]

[1]College of Information Technology, Jilin Agricultural University, Changchun 130118, China
[2]School of Data Science and Artificial Intelligence, Jilin Engineering Normal University, Changchun 130052, China
[3]Jilin Bioinformatics Research Center, Changchun 130118, China
[4]College of Engineering Technical, Jilin Agricultural University, Changchun 130118, China

Corresponding author: Shaozhong Song (e-mail: songsz@jlenu.edu.cn).

This work was supported by the Natural Science Foundation of Jilin Province (No.2020122348JC), Innovation Capacity Project on Development and Reform Commission of Jilin Province(NO.2020C019-6).We thank the anonymous reviewers for their helpful and constructive comments.

ABSTRACT Variety purity is an essential indicator in seed quality detection. Thus, it is necessary to rapidly and non-destructively detect the seed purity. Unlike traditional methods for processing hyperspectral data, this study focuses on computer vision. It aims to reconstruct RGB images from hyperspectral data and employ deep learning techniques to identify the varieties of corn seeds. Firstly， we utilized the diversity of hyperspectral band data to selectively screen the R, G, and B bands with strong feature correlations. These bands were then employed for pseudocolor reconstruction, resulting in a reconstructed dataset with distinct and more interpretive color characteristics than the original dataset. After that, we improved the classic ResNet50 model by adding the coordinate attention (CA) mechanism to the end of each residual block and replacing the global ReLU activation function with SiLU. This improvement enabled the model to capture more precise and detailed features while enhancing its predictive capability. The results showed that without improving the model, the reconstructed dataset generated by the proposed method achieved a classification accuracy of 86.28%, which was a 2.94% improvement over the original RGB dataset and it outperforms 100 randomly generated RGB combinations. While incorporating model improvements, the accuracy of the model on the reconstructed dataset reached 88.18%, surpassing other relevant models and leading to an overall accuracy improvement of 4.79%. The overall study demonstrated that combining hyperspectral image reconstruction with deep learning could be a meaningful tool for identifying and detecting the maize seed variety.

INDEX TERMS hyperspectral imaging; computer vision; variety identification; corn seeds

1. INTRODUCTION

Varietal purity plays a crucial role in seed quality testing as it directly affects seed growth and yield[1,2]. However, various maize varieties may display similar seed characteristics, and varietal mixing can occur during the growth and development process[3,4]. Visually distinguishing between the quality and variety of maize seeds is challenging, further complicating the identification of seed varieties.

Conventional methods of identifying seed varieties include morphological observation and biochemical identification[5-7]. Relying solely on morphological features like seed shape, structure, and color for identification can lead to high error rates and may be influenced by subjective judgment. Biochemical identification is an effective method for differentiating seeds with diverse genetic characteristics, but it may damage seeds and be time-consuming[8]. Therefore, conducting rapid, accurate, and non-destructive identification and screening of maize seeds before market release holds substantial potential for ensuring standardized seed markets and safeguarding seed quality[9].

Hyperspectral imaging technology combines the advantages of spectral analysis and image analysis, making it feasible for rapid and non-destructive identification of crop varieties [10]. Unlike traditional image detection methods, hyperspectral imaging techniques can accurately identify targets by examining internal components and external texture features [11-13], thus becoming one of the most effective technologies in crop quality detection.With the development and research of artificial intelligence, an increasing number of researchers are utilizing deep learning models, represented by CNN, combined with HSI to address seed quality detection issues[14,15]. Qiu[16] used a deep convolutional neural network model based on hyperspectral imaging (HSI) to classify rice seeds, achieving an accuracy rate of over 80%. Zhang[17] proposed a deep learning model combined with near-infrared hyperspectral imaging (874-1734 nm) to identify the variety of corn seeds. The extracted spectral reflection values were used to train CNN and LSTM models, and the test results showed that the model's classification accuracy rate was above 90%. Gao[18] utilized a one-dimensional convolutional neural network (1D-CNN) based on hyperspectral data (292-865 nm) to classify the presence of aflatoxin in peanuts, achieving a classification accuracy rate of 92.11% and significantly improving detection efficiency. Zhou[19] distinguished between varieties of corn seeds using near-infrared hyperspectral imaging and applied a CNN model based on partition voting for classification, achieving a classification accuracy rate of 93.33% on the test set of 6 corn varieties. Li[20] proposed a method based on deep convolutional generative adversarial networks (DCGAN) and near-infrared hyperspectral imaging technology (866.4-1701.0 nm) to identify unhealthy wheat seeds, achieving an accuracy rate of 96.67%. Deep learning algorithms have gradually become the best choice for establishing non-destructive testing models.

The current combination of deep learning algorithms and hyperspectral data primarily focuses on spectral analysis. The input data mainly consists of one-dimensional spectral information without integrating spectral and image information for training neural networks. However, relying solely on deep learning for image analysis has limitations. To address these shortcomings, we introduce a novel approach combining hyperspectral data processing using RGB pseudo-color reconstruction with deep learning techniques to classify different corn seed varieties.

Firstly, we conducted band selection to construct multiple new datasets. These datasets were then compared with the original RGB dataset in terms of accuracy to determine a reconstructed dataset with relatively higher accuracy. Building upon the usage of this dataset, improvements were made to the model to achieve further accuracy enhancement, aiming for a relatively optimal solution. The specific technical workflow is shown in Figure 1.

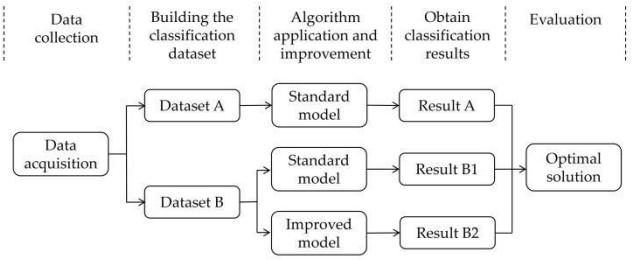

FIGURE 1. Specific technical workflow: Under the same acquisition device, dataset A refers to the conventional RGB dataset, while Dataset B refers to the reconstructed dataset obtained through RGB reconstruction. The standard model refers to the original ResNet50 model, while the improved model refers to the proposed model.

This research makes two main contributions: (1) It proposes a high-spectral data processing method based on RGB pseudo-color reconstruction, which is proven to be effective and enhances the classification performance. (2) It improves the ResNet50 model and further improves the classification performance.

The article is structured as follows: the section titled 'Materials and Methods' provides detailed descriptions of the datasets and an overview of the methods employed. This includes sequential processes such as band selection, RGB reconstruction, segmentation, data augmentation, model identification, and improvement. Subsequently, the experimental results are presented and discussed in the 'Results and Discussion' section. Finally, concluding remarks are provided in the 'Conclusions' section.

2. MATERIALS AND METHODS

2.1 Sample Preparation

Ten different varieties of corn seeds were provided by the Corn Institute of Jilin Academy of Agricultural Sciences, labeled as 001, 002n, 003n, 004n, 005n, 006n, 007n, 008n, 009n, and 010n. Experts carefully selected and certified these seeds, then manually cleaned to remove impurities and dust. All samples appeared normal, neatly arranged, and showed no visible damage. Before capturing hyperspectral images, the seeds were naturally dried and stored in a sealed container at room temperature (20-25°C). The high-definition color images are shown in Figure 2.

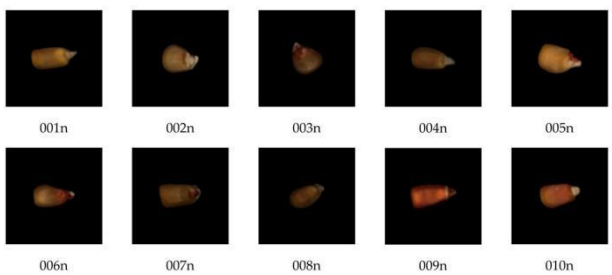

FIGURE 2. Image of maize seed varieties.

2.2 Hyperspectral Image Acquisition

Hyperspectral images of the corn seeds were captured using a line-scan hyperspectral imaging system operating in the visible/near-infrared range (392.38-1001.01 nm). Huang et al.'s paper[21] provides detailed information about the hyperspectral imaging system. To address non-uniform illumination and minimize noise in the images, we employed a black/white reference board to calibrate the experimental setup before collecting hyperspectral data. The corrected image can be obtained using Equation (1).

$$R = \frac{I - B}{W - B} \tag{1}$$

where I, B, W and R and refer to the original hyperspectral image, black reference, white reference, and calibrated image, respectively.

To acquire undistorted, clear and appropriately sized hyperspectral image data of the sample, experimental equipment parameters were repeatedly adjusted and experimental experience was summarized.After adjustments, the final experimental parameters were chosen as follows: a hyperspectral image with 462 bands in the spectral range of 392.38-1011.01 nm was acquired, the distance between the lens and the sample was set at 45 cm, the scanning speed was 0.11 cm/s, the return speed was 2.50079 cm/s, the step speed was 1.0 cm/s, the frame rate was 9.8 Hz, and the integration time was 46.701 ms.

2.3 Optimal Bands Selection

2.3.1 Principles of Image Reconstruction

The RGB image comprises three color channels: red, green, and blue. Each channel represents a single color, and different colors are formed by overlaying varying intensities[22].

In the traditional RGB color model, red occupies the highest position, green is in the middle, and blue is in the lowest position. Altering the channel order can result in different colors occupying varying positions, potentially causing color shifts and image distortion[23]. Hyperspectral data typically contains a significant amount of information in the R, G, and B bands. More diverse and accurate features can be extracted by leveraging these bands, thereby reconstructing pseudo-RGB images with distinct color characteristics. Therefore, the primary focus of our work was to select different combinations of bands to reconstruct pseudo-RGB images with diverse color features, consequently influencing recognition accuracy. An illustration of the color reconstruction effect is shown in figure 3.

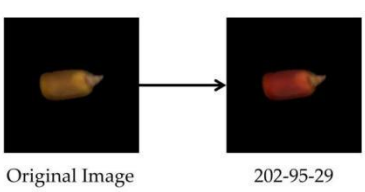

FIGURE 3. Color difference between the original image and the reconstructed image of the 202-95-29 band combination.

2.3.2 Principle and Process of Reconstruction

The correlation coefficient method is an effective approach for reducing data dimensionality. It is primarily used to select the most representative features and decrease the complexity of the dataset. Its fundamental concept lies in analyzing the correlation among features and retaining those highly correlated with the target variable. In this study, the Pearson correlation coefficient formula[24] was utilized, defined as shown in equation (2).

$$r = \frac{\mathrm{Conv}(X, Y)}{\mathrm{std}(X) \times \mathrm{std}(Y)} \tag{2}$$

where r refers to the Pearson correlation coefficient, $\mathrm{Conv}(X, Y)$ refers to the covariance between X and Y, $\mathrm{std}(X)$ and $\mathrm{std}(Y)$ refers to the standard deviations of X and Y, respectively. The Pearson correlation coefficient ranges from -1 to 1.When it is equal to 1, it indicates a perfect positive correlation.When it is equal to -1, it indicates a perfect negative correlation. When it is equal to 0, it indicates no correlation.

The correlation coefficient method effectively isolates the correlation between original variables and summarizes the dataset's information using a minimal number of variables. In contrast to Principal Component Analysis (PCA), which involves complex mathematical calculations, the correlation coefficient method offers a more intuitive and straightforward calculation process. This method directly assesses the linear correlation between features and efficiently identifies those with the highest correlation to the target variable. By calculating the correlation coefficients, it is possible to reduce the data volume and shorten the computation time. By setting a threshold or specifying the desired number of selected features, one can choose a subset of strongly correlated features that rank highly within the R (bands 173-292), G (bands 76-127), and B (bands 23-75) band intervals. These selected features are then used to compose pseudo-color images with different color characteristics.

Once we have finished selecting the bands, the reconstruction process begins. Initially, we open the original hyperspectral data and loop-read the data files to extract the three specific bands already chosen. We then store the data of these bands in a list. Next, we stack the data of the three bands along the final axis, creating a three-dimensional array. After that, we perform normalization to constrain the pixel values between 0 and 1. Finally, we multiply the normalized image data by 255 and convert it to an unsigned 8-bit integer (uint8) to transform it into an RGB image, which is saved as a JPG file. This process effectively condenses hyperspectral data, comprising hundreds of bands, into a single RGB image, ready for further analysis or visualization.

2.4 Image Preprocessing and Segmentation

After reconstructing the pseudo-RGB image, separating the maize seeds from the background is necessary. Firstly, the image is converted into a grayscale image. Then, a binary image can be obtained through automatic global

thresholding. Finally, the mask of maize seed regions can be scored by morphological filtering of the binary image and used to divide into a single maize seed, resulting in a total of 316 original images. The image segmentation process is shown in Figure 4.

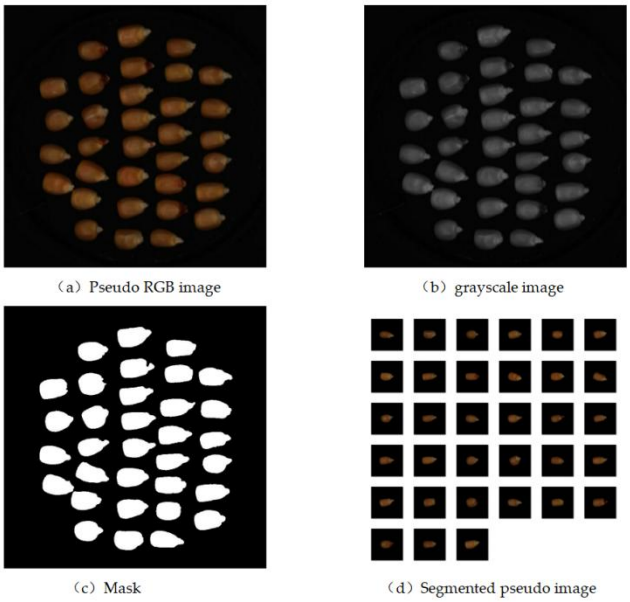

FIGURE 4. Specific technical workflow: Under the same acquisition device, dataset A refers to the conventional RGB dataset, while Dataset B refers to the reconstructed dataset obtained through RGB reconstruction. The standard model refers to the original ResNet50 model, while the improved model refers to the proposed model.

After obtaining individual maize seed images, given the limited number of samples, variations in shooting angles, and the presence of equipment noise and other interference factors, it becomes necessary to enhance the images in the dataset. This enhancement aims to learn as many unrelated patterns as possible and avoid overfitting. In data augmentation, various weather conditions are simulated for all samples in the original dataset by applying Gaussian blur, enhancing contrast by 30%, reducing contrast by 30%, increasing brightness by 30%, and decreasing brightness by 30%. Additionally, changes in shooting angles are simulated by rotating the images by 90° and 270°, horizontally flipping them, and vertically flipping them. These augmented images are then added to the original dataset. The image enhancement schematic is shown in figure 5.

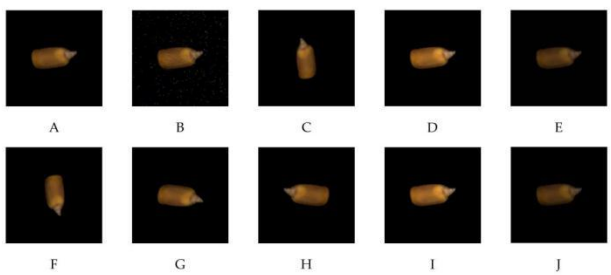

FIGURE 5. Example of image enhancement for 001n type corn seeds. (A) Original image, (B) Gaussian blur, (C) 90◦ rotation, (D) High contrast, (E) Low contrast, (F) 270◦rotation, (G) Horizontal symmetry, (H) Vertical symmetry, (I) High brightness, and (J) Low brightness.

To address many uncertainties in the experiment and because each pseudo-RGB reconstruction generates a dataset with the same number and order as the original dataset, a unified Python dataset partition script is used to divide all datasets into training and validation sets at a ratio of 8:2 for model training and validation. Due to the small sample size, we did not create a separate test set. Instead, we used the validation set as the test set to validate the results. The dataset is divided into 2504 images as the training set and 626 images as the validation set, as shown in Table 1. The final experimental results are based on the average of the results from five repetitions of the experiment.

TABLE I
DATASET PARTITION

Variety	Training Set	Validation Set	Totals
001n	240	60	300
002n	248	62	310
003n	256	64	320
004n	264	66	330
005n	240	60	300
006n	256	64	320
007n	232	58	290
008n	272	68	340
009n	240	60	330
010n	256	64	320

2.5 Building the Model

2.5.1 ResNet50 Model

In 2015, He et al. proposed ResNet[25]. Thanks to its robust feature extraction capability from pre-trained models, adaptability to the depth and complexity of small sample data, relatively fast inference speed, open-source compatibility, and good performance, ResNet50 has made an ideal choice for addressing the small sample maize seed classification problem.

2.5.2 Adding an Attention Mechanism

In recent years, attention mechanisms have been widely used in convolutional neural networks due to their ability to significantly improve network performance by refining feature maps[26-28]. The commonly used attention mechanisms include Squeeze-and-Excitation (SE)[29] and Convolutional Block Attention Module (CBAM)[30]. However, SE and CBAM do not fully consider the importance of the positional relationships and correlations between different channels. Therefore, we chose an attention method named Coordinate Attention mechanism [31], which can learn channel relationships and dependencies. The specific structure of the CA module is shown in Figure 6.

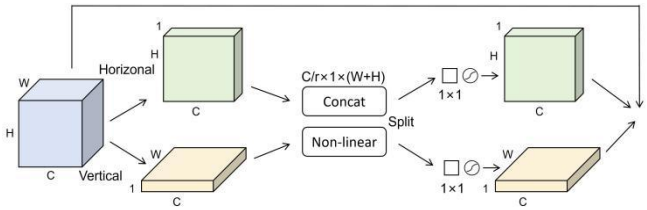

FIGURE 6. Structure of the Coordinate Attention Module.

The basic idea of the CA module is to extract the coordinate information of each position in the image and map it into a set of weights specifically for the channel dimension. This enables the coordinate information at each position to interact with other positions in the image on a global scale. Due to the unique morphological characteristics of maize seeds, the CA module captures the interdependencies within the feature maps and preserves accurate spatial information along the spatial direction. Consequently, the model can more accurately localize the target area.

2.5.3. Using SiLU Activation Function

Utilizing activation functions in neural networks boosts the model's expressive power by introducing non-linear elements. The standard ResNet50 model adopts the ReLU activation function. Figure 7 illustrates that when the input value is less than or equal to 0, the derivative of ReLU becomes 0. Consequently, neurons may become 'dead,' impeding weight updates and potentially leading to information loss.

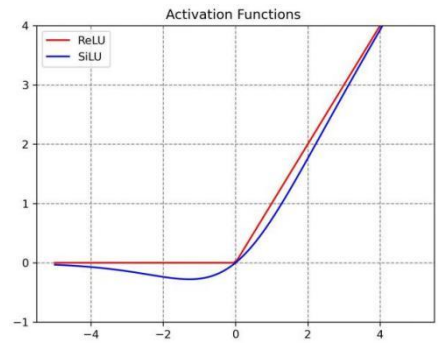

FIGURE 7. Relu and Silu activation function curves.

To tackle this concern, we opted to substitute the ReLU activation function with SiLU. SiLU offers numerous advantages over traditional activation functions: unbounded from above, bounded from below, smooth, and non-monotonic[32]. The SiLU activation function outperforms the ReLU activation function in deep models while retaining the advantages of ReLU, namely the ability to mitigate the gradient vanishing problem and addressing the aforementioned drawbacks of the ReLU function.

In summary, we have incorporated the SiLU activation function into the maize seed recognition model to enhance prediction capability.

2.5.4. Proposed Methods

To capture cross-channel information and incorporate spatial information for more accurate localization of target regions, enhancing the understanding of inputs and feature extraction capabilities, we designed our model by integrating the ResNet50 module with the CA module. As shown in Figure 8, we integrated the CA module into the tail of each residual block in the Conv5_x of ResNet50 to enhance feature representation. Additionally, we replaced the global ReLU activation function with the SiLU activation function, which partially solves the issue of weights not being updated when inputs are less than or equal to zero and the inability to process negative inputs directly. This further improves the recognition performance of the model.

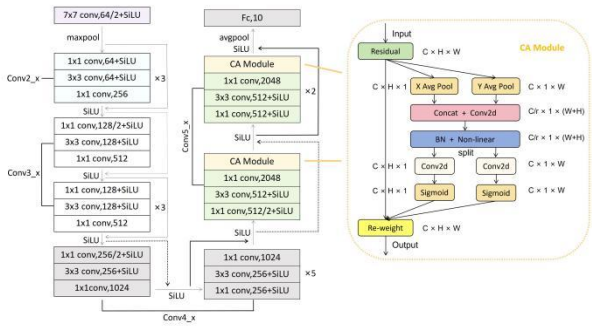

FIGURE 8. Structure of the Coordinate Attention Module.

2.6 Evaluation Indicators

This article utilizes Accuracy, Precision, Recall, and F1-score as performance evaluation indicators for the model. Accuracy refers to the proportion of correctly predicted instances among the total number of predictions. Precision refers to the probability of a specific category being correctly predicted among all predicted results. Recall refers to the probability of a specific category being correctly predicted among all actual values. The F1 score is the harmonic mean of precision and recall.

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

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

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

$$\text{F1} = 2 \times \frac{(\text{Precision} \times \text{Recall})}{(\text{Precision} + \text{Recall})} \tag{6}$$

where TP refers to the correctly classified positive samples, FP refers to the negative samples mistakenly classified as positive, TN refers to the correctly classified negative samples, and FN refers to the positive samples mistakenly classified as negative.

3. RESULTS AND DISCUSSIONS

3.1 Experimental Setup

We implemented our approach based on PyTorch. The processor of the experimental workstation was Xeon 5220R; the GPU was an NVIDIA TESLA T4; the running memory was 128 GB RAM; the operating system is Windows 10. The

software experimental configuration environment was based on Python 3.8.16; Pytorch 1.13.1; and CUDA 11.6. The specific parameter settings in the experiment are shown in Table 2.

TABLE II
TRAINING HYPERPARAMETER INFORMATION

Parameter	Value or Name
Training epochs	50
Batch size	16
Learning rate	0.0001
Weight decay	0.0001
Optimizer	Adam

Due to the limited number of samples, we employ transfer learning methods to enhance the accuracy and generalization of our model. We initialize a classification model by loading pre-trained weights from the ImageNet dataset. Subsequently, we continuously fine-tune it through training.

3.2 Image Reconstruction Validation

To validate the effectiveness of the proposed ideas and corresponding strategies, we selected the top three sets of more informative R, G, and B bands for verification. The selected R bands are the 173rd, 179th, and 181st bands, the G bands are the 123rd, 124th, and 127th bands, and the B bands are the 71st, 73rd, and 74th bands. These three bands were combined to form 27 pseudo-RGB images for recognition and compared with the original RGB dataset. The results, as shown in Figure 9, revealed an accuracy rate of 83.39% for the original RGB dataset. Out of the 27 combinations of bands, only two had lower accuracy rates than the original image, while the remaining 25 combinations showed higher accuracy rates. Specifically, the combination of bands 179-127-71 achieved the highest accuracy rate at 86.28% within the range, representing a 2.89% improvement over the original image. This indicates that the selected band features contain more valuable information, providing a more detailed and comprehensive description, thus enhancing the interpretability and predictive capability of the features.

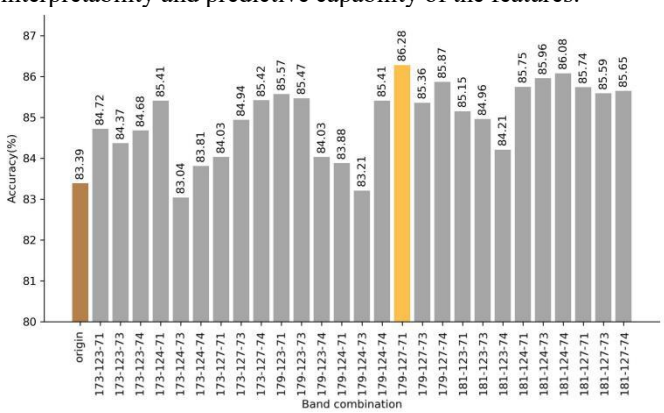

FIGURE 9. Results of the comparative experiment on 27 band combinations.

In addition, we randomly generated 100 band combinations within the R, G, and B spectral ranges for comparison. Out of the 100 randomly generated combinations, only 45 achieved higher accuracy than the original image. The accuracy of the top 20 band combinations is illustrated in Figure 10. It can be observed from the figure that the band combination of 176-123-56 achieved the highest accuracy within the range, reaching 85.85%. This represents a 2.46% improvement compared to the original image. However, it is still lower than the accuracy of the previously selected band combination of 179-127-71. This indicates that the proposed strategy can more effectively screen out more important feature bands, demonstrating the feasibility and efficiency of the proposed approach.

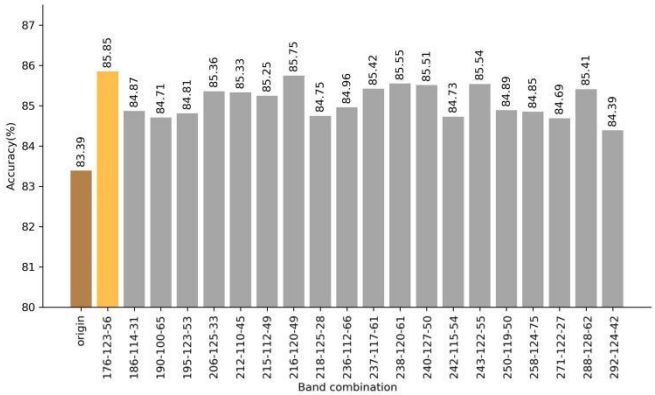

FIGURE 10. Results of the comparative experiment on 100 randomly generated band combinations (top 20)

3.3 Model Improvement Validation

3.3.1 Comparison Experiments of Different Models

To validate the effectiveness and state-of-the-art performance of the new network model, we conducted a comparison with five classical convolutional neural networks, namely ResNet50 (original model), ResNext50, Res2Net50, EfficientNet, and Convnext_T. This comparison was conducted using the dataset with the band combination of 179-127-71. The results are shown in Table 3, Figures 11 and 12.

TABLE III
COMPARISON EXPERIMENTS OF DIFFERENT MODELS

Model	Accuracy	Loss	Average Precision	Average Recall	Average F1-score
ResNet50	86.28%	0.5276	87.21%	86.44%	0.8682
ResNext50	86.79%	0.5256	87.47%	87.08%	0.8727
Res2Net50	85.75%	0.5454	86.84%	85.95%	0.8639
EfficientNet	86.93%	0.4652	87.74%	87.11%	0.8742
Convnext_T	84.12%	0.6076	85.68%	84.38%	0.8503
Our model	88.18%	0.4438	89.09%	88.21%	0.8865

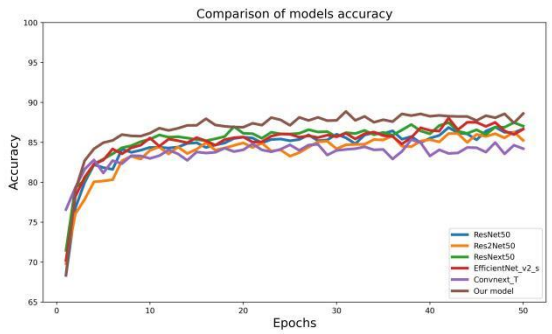

FIGURE 11. Results of the accuracy of different model comparison experiments.

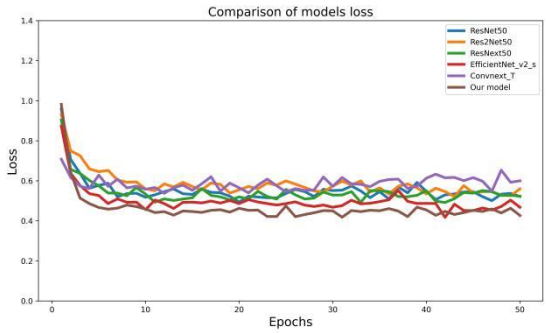

FIGURE 12. Results of the loss of different model comparison experiments.

The comparative experiments have shown that the proposed network model in this paper achieved recognition accuracy of 88.18%, with a loss of 0.4438. The average precision is 89.09%, the average recall is 88.21%, and the average F1 score is 0.8825. All evaluation metrics of this model surpass those of other models; this proves the model's superiority and provides a valuable reference for its application in identifying corn seed varieties in the future.

3.3.2. Ablation Experiments

We conducted an ablation experiment to investigate the impact of the CA module on the classification performance of ResNet50. As shown in Table 4, the results indicate that adding an additional CA module does not improve the model's performance. In contrast, optimal results are achieved only when the CA module is added at a specific position: the end of each residual block in Conv5_x. This suggests that incorporating the CA module at the end of subsequent residual blocks in the model enables better adaptation to the correlation and importance of local features, thereby enhancing attention and capturing important features more effectively.

TABLE Ⅳ
COMPARATIVE RESULTS OF INCORPORATING CA ATTENTION MECHANISM AT DIFFERENT POSITIONS. "√" INDICATES THE INCLUSION OF CA ATTENTION MECHANISM AT THAT POSITION.

Num	Head	Conv2_x	Conv3_x	Con4_x	Conv5_x	Tail	Acc
1							86.28%
2	√						85.81%
3		√	√	√	√		85.68%
4			√	√	√		86.51%
5				√	√		87.14%
6					√		87.64%
7						√	87.08%

In our experiment, we found that adding different attention mechanisms also impacted the model's performance when all other factors were kept constant. As shown in Table 5, the CA module has a more significant impact on improving the performance of the network model. This suggests that, compared to other attention mechanisms, the CA module possesses advantages in fine-grained focus and spatial adaptability, allowing for better capture of local features while reducing the influence of background noise, thereby enhancing the model's accuracy.

TABLE Ⅴ
COMPARISON OF THE RESULTS OBTAINED WITH DIFFERENT ATTENTION MECHANISMS

Num	Model	Insert Position	Attention	Accuracy
1	ResNet50	None	None	86.28%
2	ResNet50	Conv5_x	CA	87.64%
3	ResNet50	Conv5_x	CBAM	86.62%
4	ResNet50	Conv5_x	DAN	85.95%
5	ResNet50	Conv5_x	ECA	87.13%
6	ResNet50	Conv5_x	SE	86.85%

The choice of activation function plays a crucial role in the training process, significantly impacting the model's performance. In our experiment, after incorporating the CA module, we substituted the original ReLU activation function with other activation functions to evaluate the model's performance. As shown in Table 6, the SiLU activation function achieves the highest accuracy. Two factors contribute to this outcome: 1) The SiLU activation function maintains a non-zero mean within the input range, which preserves more information and enhances the model's expressive capability. 2) With a smoother curve and continuous derivatives across various positions, the SiLU activation function facilitates better convergence of the optimization algorithm and mitigates gradient explosion and vanishing issues.

TABLEⅥ
EXPERIMENTAL COMPARISON RESULTS OF DIFFERENT ACTIVATION FUNCTIONS

Num	Model	Activate Function	Accuracy
1	ResNet50+CA	ReLU(origin)	87.64%
2	ResNet50+CA	ELU	87.19%
3	ResNet50+CA	LeakyReLU	87.96%
4	ResNet50+CA	PRelU	87.44%
5	ResNet50+CA	SilU	88.18%

To verify the recognition capability of the proposed method combined with the improved model, the visualization of the confusion matrix is shown in Figure 13. In the original dataset and model, 10, 4, and 3 samples of the 10th class seed were misclassified as the 1st, 5th, and 6th classes, respectively. In contrast, with the improved dataset and network model, only 5 sample of the 10th class seed was misclassified as the 1th class, without misclassifying it as the 5st or 6th class seed. Moreover, the proposed approach effectively reduced the misclassification of the 6th class seed as the 5th class, the misclassification of the 7th class seed as the 2nd class, and the misclassification of the 8th class seed as the 7th class.

In summary, the method proposed in this study and the enhanced model substantially decreased the initial error rate. It bolstered the model's accuracy, enhancing its efficacy in identifying various corn seed varieties. Furthermore, a detailed analysis utilizing the confusion matrix of the improved network model unveiled a notable absence of misclassifications among the 5th class seed samples. However, significant errors still existed in identifying the 3rd and 4th class samples.

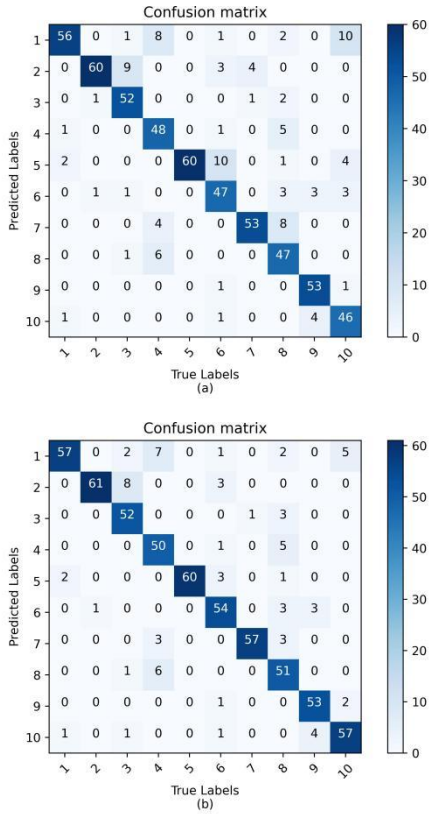

FIGURE 13. Visualization of the model's confusion matrix. (a) Confusion matrix visualization using the original dataset and the original model. (b) Confusion matrix visualization using the reconstructed dataset and the improved model.

3.4 Comparison of results with traditional methods

To compare with traditional hyperspectral processing classification methods, we sequentially applied Min-Max Scaling (MMS), Moving Average Smoothing (MA), Linear Discriminant Analysis (LDA), and the machine learning algorithm GaussianNB to the dataset. These techniques were used for normalization, noise reduction, dimensionality reduction, and recognition, respectively, to construct conventional recognition methods. As shown in the table 7, it can be observed that compared to traditional methods, our approach achieves higher values in Accuracy, Precision, Recall, and F1-score. This indicates the superiority of the proposed method, enabling more precise feature extraction and thus improving the accuracy of maize seed recognition.

TABLEⅦ
COMPARISON BETWEEN TRADITIONAL METHOD AND OUR METHOD

Method	Accuracy	Average Precision	Average Recall	Average F1-score
MMS-MA-LDA-GaussianNB	79.81%	83.63%	83.04%	0.8333
RGB reconstruction-Our model	88.18%	89.09%	88.21%	0.8865

3.5 Comparison of the Experimental Model and Expert Identification Results

The accuracy comparison between expert identification and the method proposed by us is shown in Table 8. In the maize seed variety classification task, our method has achieved good results in identifying various types of maize seeds. Compared with the results of expert identification, our method has shown varying degrees of improvement in accuracy across ten categories. This demonstrates that our method has a significant advantage in maize seed variety classification accuracy. It can better capture the features of seed images and conduct effective classification.

TABLEⅧ
COMPARISON OF EXPERT EVALUATION AND ACCURACY OF OUR METHOD

Level	Expert Identification Accuracy	Our Method Accuracy
001n	83.33	95%
002n	88.71	98.39%
003n	71.88	81.25%
004n	60.61	75.76%
005n	93.33	100%
006n	78.13	84.38%
007n	87.93	98.28%
008n	66.18	75%
009n	81.67	88.33%
010n	78.13	89.06%

This experiment's high precision in classifying corn seed varieties underscores its practical importance. By combining hyperspectral image reconstruction with model enhancements, various maize seed types can be accurately classified. This approach mitigates the subjectivity and human biases inherent in relying solely on expert identification. This demonstrates the proposed approach's necessity and efficacy in addressing the maize seed variety classification challenge, providing a precise and efficient solution for such tasks.

4. CONCLUSION

This study proposes a processing method that combines hyperspectral RGB band image reconstruction with deep learning. Specifically, it selects the R, G, and B feature bands with more prominent color characteristics from the diverse hyperspectral bands to reconstruct a new dataset. This reconstructed dataset is then used for recognition and comparison using the classical deep learning model ResNet50. This paper evaluated 27 feature band combinations selected using the correlation coefficient method. Among them, 25 combinations achieved higher accuracy rates than the original dataset. The highest accuracy rate improved by 2.89% compared to the original image, surpassing the accuracy rates of all randomly generated 100 band combinations. These results confirm the effectiveness and efficiency of the proposed method and the band selection strategy. In this study, we improved ResNet50 by adding a CA module to the subsequent residual blocks to capture the features of maize seeds better.Additionally, we replaced the traditional ReLU activation function with the SiLU activation function to enhance the model's expressive power. Our proposed model has more advantages by comparing five commonly used network models for image classification: ResNet50, ResNext50, Res2Net50, EfficientNet, and

Convnext_T. It achieved an accuracy of 88.18% on the reconstructed dataset, which is a total improvement of 4.79% in overall accuracy compared to the original data and model.

Due to the current limited data conditions and the excessively large number of possible band combinations, we can only strive to find relatively effective band selection methods. In the future, we plan to supplement our sample size with other types of maize seeds, such as defective seeds, and explore more effective band selection methods to choose reconstruction combinations with better results, thus improving identification accuracy. In addition, we plan to continue optimizing the model to achieve faster and more accurate classification of corn seed images.

REFERENCES

[1] Ambrose, A.; Kandpal, L.M.; Kim, M.S.; Lee, W.; Cho, B. High speed measurement of corn seed viability using hyperspectral imaging. Infrared Phys. Technol. 2016, 75, 173-179, doi: https://doi.org/10.1016/j.infrared.2015.12.008.

[2] Ma, T.; Tsuchikawa, S.; Inagaki, T. Rapid and non-destructive seed viability prediction using near-infrared hyperspectral imaging coupled with a deep learning approach. Comput. Electron. Agric. 2020, 177, 105683.

[3] Zhang, X.; Liu, F.; He, Y.; Li, X. Application of hyperspectral imaging and chemometric calibrations for variety discrimina-tion of maize seeds. Sensors-Basel 12 2012, 17234-17246.

[4] Yang, X.; Hong, H.; You, Z.; Cheng, F. Spectral and image inte-grated analysis of hyperspectral data for waxy corn seed variety classification. Sensors-Basel 15 2015, 15578-15594.

[5] Ballard, T.O.; Foley, M.E.; Bauman, T.T. Germination, viability, and protein changes during cold stratification of giant ragweed (Ambrosia trifida l.) seed. J. Plant Physiol. 1996, 149, 229-232, doi: https://doi.org/10.1016/S0176-1617(96)80201-X.

[6] Rao, P.S.; Bharathi, M.; Reddy, K.B.; Keshavulu, K.; Rao, L.; Neeraja, C.N. Varietal identification in rice (Oryza sativa) through chemical tests and gel electrophoresis of soluble seed proteins. Indian J. Agric. Sci. 2012, 82, 304-311.

[7] Boniecka, J.; Kotowicz, K.; Skrzypek, E.; Dziurka, K.; Rewers, M.; Jedrzejczyk, I.; Wilmowicz, E.; Berdychowska, J.; Dąbrowska, G.B. Potential biochemical, genetic and molecular markers of deterioration advancement in seeds of oilseed rape (Brassica napus L.). Ind. Crop. Prod. 2019, 130, 478-490, doi: https://doi.org/10.1016/j.indcrop.2018.12.098.

[8] Cui, Y.J.; Xu, L.J.; An, D.; Liu, Z.; Gu, J.C.; Li, S.M.; Zhang, X.D.; Zhu, D.H. Identification of maize seed varieties based on near infrared reflectance spectroscopy and chemometrics. Int. J. Agric. Biol. Eng. 2018, 11, 177-183.

[9] Tenaillon, M.I.; Charcosset, A. A European perspective on maize history. C. R. Biol. 2011, 334, 221-228.

[10] Wu, Z.; Li, G.; Yang, R.; Fu, L.; Li, R.; Wang, S. Coefficient of restitution of kiwifruit without external interference. J. Food Eng. 2022, 327, 111060, doi: https://doi.org/10.1016/j.jfoodeng.2022.111060.

[11] Wei, Y.; Wu, F.; Xu, J.; Sha, J.; Zhao, Z.; He, Y.; Li, X. Visual detection of the moisture content of tea leaves with hyperspectral imaging technology. J. Food Eng. 2019, 248, 89-96, doi: https://doi.org/10.1016/j.jfoodeng.2019.01.004.

[12] Xuan, W.; Jin-Cheng, H.; Da-Peng, Y.; Deng-Fei, J. Navel Orange Maturity Classification by Multispectral Indexes Based on Hyperspectral Diffuse Transmittance Imaging. J. Food Qual. 2017, 2017, 1-7.

[13] Sun, J.; Cong, S.; Mao, H.; Wu, X.; Yang, N. Quantitative detection of mixed pesticide residue of lettuce leaves based on hyperspectral technique. J. Food Process Eng. 2018, 41, -.

[14] Elmasry, G.; Mandour, N.; Al-Rejaie, S.; Belin, E.; Rousseau, D. Recent Applications of Multispectral Imaging in Seed Phenotyping and Quality Monitoring—An Overview. Sensors 2019, 19, 1090.

[15] Gao, T.; Chandran, A.K.N.; Paul, P.; Yu, H. HyperSeed: An End-to-End Method to Process Hyperspectral Images of Seeds. Sensors 2021.

[16] Qiu, Z.; Chen, J.; Zhao, Y.; Zhu, S.; He, Y.; Zhang, C. Variety Identification of Single Rice Seed Using Hyperspectral Imaging Combined with Convolutional Neural Network. In Applied Sciences, 2018; Volume 8.

[17] Zhang, C.; Zhao, Y.; Yan, T.; Bai, X.; Xiao, Q.; Gao, P.; Li, M.; Huang, W.; Bao, Y.; He, Y.; et al. Application of near-infrared hyperspectral imaging for variety identification of coated maize kernels with deep learning. Infrared Phys. Technol. 2020, 111, 103550, doi: https://doi.org/10.1016/j.infrared.2020.103550.

[18] Gao, J.; Zhao, L.; Li, J.; Deng, L.; Ni, J.; Han, Z. Aflatoxin rapid detection based on hyperspectral with 1D-convolution neural network in the pixel level. Food Chem. 2021, 360, 129968, doi: https://doi.org/10.1016/j.foodchem.2021.129968.

[19] Zhou, Q.; Huang, W.; Tian, X.; Yang, Y.; Liang, D. Identification of the variety of maize seeds based on hyperspectral images coupled with convolutional neural networks and subregional voting. J. Sci. Food Agric. 2021, 101, 4532-4542.

[20] Li, H.; Zhang, L.; Sun, H.; Rao, Z.; Ji, H. Discrimination of unsound wheat kernels based on deep convolutional generative adversarial network and near-infrared hyperspectral imaging technology. Spectrochimica Acta Part a: Molecular and Biomolecular Spectroscopy 2022, 268, 120722, doi: https://doi.org/10.1016/j.saa.2021.120722.

[21] Huang, M.; Tang, J.; Yang, B.; Zhu, Q. Classification of maize seeds of different years based on hyperspectral imaging and model updating. Comput. Electron. Agric. 2016, 122, 139-145, doi: https://doi.org/10.1016/j.compag.2016.01.029.

[22] Hunt, R.W.G. The reproduction of colour; The reproduction of colour, 2004.

[23] Qiuqi, R. Digital Image Processing [M]; Electronic Industry Press, 2001.

[24] Pearson, K. VII. Mathematical contributions to the theory of evolution. — III. Regression, heredity, and panmixia. Philosophical Transactions of the Royal Society of London. Series a, Containing Papers of a Mathematical Or Physical Character 1896, 253-318.

[25] He, K.; Zhang, X.; Ren, S.; Sun, J. Deep Residual Learning for Image Recognition. Ieee 2016.

[26] Zhu, X.; Cheng, D.; Zhang, Z.; Lin, S.; Dai, J. An Empirical Study of Spatial Attention Mechanisms in Deep Networks. 2019.

[27] Yan, H.W.; Liu, Z.Y.; Cui, Q.L.; Hu, Z.W. Multi-target detection based on feature pyramid attention and deep convolution network for pigs. Transactions of the Csae 2020, 36, 193-202.

[28] Choi, H.; Cho, K.; Bengio, Y. Fine-grained attention mechanism for neural machine translation. Neurocomputing 2018, 284, 171-176, doi: https://doi.org/10.1016/j.neucom.2018.01.007.

[29] J., H.; L., S.; S., A.; G., S.; E., W. Squeeze-and-Excitation Networks. Ieee Trans. Pattern Anal. Mach. Intell. 2020, 42, 2011-2023, doi: 10.1109/TPAMI.2019.2913372.

[30] Woo, S.; Park, J.; Lee, J.Y.; Kweon, I.S. CBAM: Convolutional Block Attention Module. Springer, Cham 2018.

[31] Hou, Q.; Zhou, D.; Feng, J. Coordinate attention for efficient mobile network design. In Proceedings of the IEEE/CVF conference on computer vision and pattern recognition, 2021; pp. 13713-13722.

[32] Ramachandran, P.; Zoph, B.; Le, Q.V. Searching for Activation Functions, 2017. substrates," *IEEE Photon. J.*, vol. 5, no. 2, Apr. 2013, Art. no. 2600111

[33] Azodolmolky *et al.*, Experimental demonstration of an impairment aware network planning and operation tool for transparent/translucent optical networks," *J. Lightw. Technol.*, vol. 29, no. 4, pp. 439–448, Sep. 2011.

Jian Li received the Ph.D degrees from the Jilin University, Changchun, China, in 2013.

He is now working as a professor at the School of Information Technology, Jilin Agricultural University, and his main research interests are artificial intelligence and bioinformatics

E-mail: liemperor@163.com

Fan Xu received the B.E. degree from Jilin Agricultural University in Changchun ,China,in 2022,

He is currently studying for a Masters in Computer and Science at Jilin Agricultural University, Mainly engaged in the research of computer vision image classification

E-mail: 20221128@mails.jlau.edu.cn

Shaozhong Song received the Ph.D degrees from the Jilin University, Changchun, China, in 2012.

He is now working as a professor at the School of Data Science and Artificial Intelligence, Jilin Engineering Normal University, and his main research interests are digital agriculture, Pattern recognition and computer application technology

E-mail: songsz@jlenu.edu.cn

Ji Qi received the doctor‘s degree from Dalian Maritime University, P.R. China.

He is now working at College of Engineering Technical, Jilin Agricultural University. His research interest include information engineering and control, intelligent monitoring and algorithm and information and technology science.

E-mail: jiq@jlau.edu.cn

Junling Liu received the Ph.D degrees from the Jilin University, Changchun, China, in 2014.

She is now working as a professor at the School of Data Science and Artificial Intelligence, Jilin Engineering Normal University, and her main research interests are computer vision and network security

E-mail: 187009182@qq.com