

RESEARCH ARTICLE

Intelligent classification of flue-cured tobacco based on improved network model of six channel ResNet

Shiyu Wang¹, Runjie Zhuang¹, Wanping Lv¹, Tianzhen Lin¹, Yongxian Wen^{1,2,*}

¹College of Computer and information Science, ²Institute of Statistics and Applications, Fujian Agriculture and Forestry University, Fuzhou, Fujian, China.

Received: August 15, 2024; accepted: December 31, 2024.

Flue-cured tobacco is the most important raw material for cigarette production, and efficient and accurate intelligent grading of flue-cured tobacco leaves is of great significance to the acquisition and correct use of tobacco. Currently, there are few studies on the classification of flue-cured tobacco by using both front and back images of tobacco leaves, while the back image of flue-cured tobacco leaves has enriched information. To harness the back information of these leaves and enhance grading accuracy, a revolution ResNet network model, Evo-ResNet, was proposed in this study based on front and back images of flue-cured tobacco leaves. Evo-ResNet had six channel with added Squeeze-and-Excitation (SE) attention mechanism. Six grades pictures of Cuibi 1 and Yunyan 87 tobacco species were collected using mobile phone, respectively. Evo-ResNet was compared with three-channel models of ResNet, GoogLeNet, VGGNet, and AlexNet with only front images of flue-cured tobacco leaves and both front and back images of flue-cured tobacco leaves. The results showed that the loss of train set and the accuracy of validation set had better results while the front and back pictures of flue-cured tobacco leaves were used meantime. For the test set, the accuracy, recall, precision, and F1-score of Evo-ResNet were better than that of other three-channel network models, expect for GoogleNet. The accuracy, recall, precision, and F1-score value of Cuibi 1 were 95.83%, 95.83%, 96.16%, and 95.78%, respectively, while those of Yunyan 87 were 99.07%, 99.07%, 99.10% and 99.08%, respectively. There was better generalization ability for Evo-ResNet. Because of the convenience for obtaining image of flue-cured tobacco by mobile phone, the proposed Evo-ResNet network model had better applicability in practice, which provided a new idea for the application of intelligent classification of flue-cured tobacco leaves in purchasing practice of flue-cured tobacco.

Keywords: tobacco leaf grading; Evo-ResNet model; six channels; Se module.

*Corresponding author: Yongxian Wen, College of Computer and information Science, Fujian Agriculture and Forestry University, Fuzhou 350002, Fujian, China. Email: Wen9681@sina.com.

Introduction

The tobacco industry holds significant prominence in China. Flue-cured tobacco is a critical raw material in cigarette production, where its quality significantly influences its pricing. Therefore, rational grading of flue-cured tobacco is crucial for maximizing economic

benefits. Tobacco grading is based on the quality of flue-cured tobacco. However, in most tobacco production areas, manual grading methods are still predominant presently. Tobacco grading workers assess flue-cured tobacco through visual, tactile, and olfactory observations. Despite standardized rigorous professional training, individual differences persist in grading

efficiency and accuracy among workers [1, 2]. During the grading process, workers need to be careful with each piece of flue-cured tobacco. Prolonged grading sessions can induce visual fatigue among workers, thereby impacting both the efficiency and accuracy of the grading process. This situation may lead to disputes between tobacco farmers and purchasing stations regarding the grading outcomes. In response to this challenge, the widespread adoption of computer technology has emerged as a common solution to enhance tobacco grading practices.

It is highly feasible to apply image recognition techniques such as machine vision and convolutional neural networks, commonly used in image classification, to grading flue-cured tobacco. Intelligent classification of tobacco dates back to 1997 when Korean scholars first proposed the use of machine vision methods for this purpose [3]. Presently, intelligent grading encompasses four main approaches including spectral technology-based classification [4, 5], fuzzy mathematics-based classification [6, 7], machine vision methods [8, 9], and classification utilizing convolutional neural networks. In recent years, the organization of the ImageNet Large Scale Visual Recognition Challenge (ILSVRC) has significantly accelerated the development of convolutional neural networks (CNNs) such as visual geometry group (VGGNet) and residual neural network (ResNet). Researchers have meticulously refined these foundational CNN architectures and achieved impressive results when applied to tobacco classification. Liang *et al.* introduced a tobacco grading approach leveraging machine vision, which employed CNNs including ResNet, AlexNet, VGGNet as foundational learners and a support vector machines (SVM) classifier as the meta-learner [10]. The stacking model fusion strategy significantly enhanced accuracy over individual models as evidenced by comparative analysis. Wang *et al.* achieved 95.23% accuracy in tobacco leaf classification *via* InceptionV3 with transfer learning for training and integration of the extreme point jumping algorithm [11]. Lu *et al.*

proposed a deep learning approach for tobacco leaf classification, which incorporated multi-scale feature fusion [12]. Specifically, the method utilized ResNet50 for feature extraction and incorporated a compression excitation module to assign different feature weights to different channels. Subsequently, these features were fused with a feature pyramid to realize multi-scale feature representation of tobacco leaves, resulting in 5.21% higher accuracy than classical CNNs. Other researchers compared the performance of four classical CNNs pretrained on Imagenet for tobacco grading and found that VGG-16 performed effectively on Anhui Wannan Tobacco data. Subsequently, they evaluated the impact of five different optimizers and concluded that the Adagrad optimizer achieved the highest overall performance with an accuracy rate of 96% [13]. Additionally, building upon ResNext50, researchers introduced the CA attention mechanism [14], which enhanced the channel attention to features and thereby improved the grading accuracy of the enhanced model by 9.4% higher than the base ResNet50 network [15].

Currently, most scholars and grading experts only rely on the front images of tobacco leaves, employing visual, tactile, and olfactory assessments for flue-cured tobacco classification, while the utilization of back side data remains restricted [12]. Nevertheless, the back images of tobacco leaves, which are richer in texture and structural complexity than the front, could enhance grading accuracy if fully utilized. Therefore, based on the ResNet convolutional neural network, this research proposed an improved algorithm that aligned and integrated both front and back images data of tobacco leaves into a six-channel tensor, which adjusted the input channel of the initial convolutional layer in ResNet network model to six and added the squeeze-and-excitation (SE) attention mechanism, for intelligent classification of flue-cured tobacco. This study had significant implications for both the current tobacco industry and the research community, which achieved a more convenient, efficient, and accurate intelligent classification in the tobacco

Table1. Grade quality regulations for flue-cured tobacco.

Groups	Code	Maturity	Leaf structure	Body	Oil	Color intensity	Length (cm)	Waste (%)	
X	F	X2F	ripe	open	less thin	less oily	moderate	35	25
		C2F	ripe	open	medium	oily	strong	40	15
C	F	C3F	ripe	open	medium	oily	moderate	35	25
		C4F	ripe	open	less thin	less oily	moderate	35	30
B	F	B2F	ripe	firm	fleshy	oily	strong	40	20
		B3F	ripe	close	fleshy	oily	moderate	35	30

Notes: B, C, and X represented upper, middle, and lower part of tobacco leaves, respectively. F represented orange tobacco.

industry by using mobile phones of any model and pixel to get the images. Further, this study extended the application of machine vision and deep learning in agriculture, particularly in tobacco classification. The utilization of back images of flue-cured tobacco leaves offered new ideas for advancements in tobacco leaf classification intelligence and enriched the current theoretical framework of tobacco classification, which not only served as a reference for the intelligent classification of other crops but also facilitated a more efficient production model and scientific resource allocation, thereby positively contributing to the modernization of the tobacco industry.

Materials and methods

Resource of tobacco leaves

Six grades of cured tobacco including X2F, C2F, C3F, C4F, B2F, B3F sourced from Fujian province purchasing stations (Fujian, China) were employed in this study. According to the National Standard of the People's Republic of China: Flue-cured Tobacco, GB2635-92 [16], the tobacco grades were designated by the part of the plant where the leaves were collected with the upper part as B, middle as C, lower as X, while the quality grade was expressed by Arabic numerals with lower values indicating higher quality, and color with F for orange. The determination of the grade of tobacco included the maturity of flue-cured tobacco, leaf structure, body, oil, color intensity, length reaching a certain level of regulation when the waste does not exceed the permissible degree of a certain level. The specific criteria

were shown in Table 1. Two flue-cured tobacco varieties, Cuibi 1 and Yunyan 87, were analyzed in this study with C2F, C3F, C4F grades from Cuibi 1 procured from purchasing stations in Jiangle County and the same grades of Yunyan 87 from purchasing stations in Taining County, both located in Sanming, Fujian, China. Furthermore, X2F, B3F, and B2F grades of Cuibi 1 and Yunyan 87 were purchased from the stations in Yong'an, Fujian, China. Before sampling, the tobacco leaves were authenticated by grading experts from the Fujian Tobacco Company Sanming Branch (Sanming, Fujian, China).

Collection of tobacco leaf images

To reduce the model's dependency on shooting conditions and improve generalization, diverse devices were used to capture flue-cured tobacco leaf data under different lighting conditions at different times. All original images were acquired from the paper by Huang *et al.* and were collected in August and November 2023 [17]. In August, the images of tobacco leaves for C2F, C3F, C4F grades were captured under both indoor and outdoor natural lighting conditions at different times of the day (morning, afternoon, evening). In November, the images for X2F, B2F, B3F grades were taken under outdoor natural lighting conditions at the same times as that in August. The image capture devices included IQOO Neo 855 (IQOO, Dongguan, Guangdong, China), Huawei Nova 5i Pro (Huawei Technologies Co., Ltd., Shenzhen, Guangdong, China), Realme X50 Pro (Realme, Dongguan, Guangdong, China), and Vivo X60t Pro (Vivo, Dongguan, Guangdong, China). The different conditions were designed to enhance the

Table 2. The modeling pictures data set of flue-cured tobacco leaves for Cuibi 1 and Yunyan 87.

Species		Cuibi 1						Yunyan 87					
Data set		Grades						Grades					
		X2F	C2F	C3F	C4F	B2F	B3F	X2F	C2F	C3F	C4F	B2F	B3F
Training set	Front	288	288	288	288	288	288	432	432	432	432	432	432
	Back	288	288	288	288	288	288	432	432	432	432	432	432
Total		576	576	576	576	576	576	864	864	864	864	864	864
Validation set	Front	36	36	36	36	36	36	54	54	54	54	54	54
	Back	36	36	36	36	36	36	54	54	54	54	54	54
Total		72	72	72	72	72	72	108	108	108	108	108	108
Test set	Front	36	36	36	36	36	36	54	54	54	54	54	54
	Back	36	36	36	36	36	36	54	54	54	54	54	54
Total		72	72	72	72	72	72	108	108	108	108	108	108

model's generalization capability [17].

Construction of the dataset

After taking flue-cured tobacco pictures, preliminary data organization was important to effectively master the characteristics of the flue-cured tobacco. It was necessary to ensure that the front and back sides of the flue-cured tobacco pictures were fully unfolded and the pictures with curled flue-cured tobacco leaves were deleted to ensure the accuracy and reliability of the data. It was also important to verify the one-to-one correspondence between front and back images and to eliminate redundant data that were missed or duplicated during the shooting process to ensure data consistency. A total of 360 front and 360 back images for each grade of Cuibi 1 were selected, resulting in a total of 720 images per grade, which ended in a total of 4,320 images of Cuibi 1 from the organized data. For each grade of Yunyan 87, 540 front and 540 back images were selected with a total of 1,080 images per grade and a total of 6,480 images for Yunyan 87. The selected image data were divided into training, validation, and test sets in an 8:1:1 ratio with the detailed selected images dataset for grading purpose shown in Table 2. The image data in the training set underwent random rotation to enhance the model's ability to learn more diverse features and achieve a higher semantic representation level. The image resolution was standardized to 224×224 and converted to tensor format, enabling efficient

GPU-based training. Tensor splicing combined front and back images data into six channels as required by the algorithm in this study before model input. Data normalization was subsequently applied.

Evolution ResNet (Evo-ResNet)

This study proposed an enhanced conventional ResNet network model named evolution ResNet (Evo-ResNet) with the algorithm's flowchart shown in Figure 1.

(1) Modification of data folder structure

Both front and back images of tobacco leaves were input into the model on a one-to-one correspondence, enabling model to learn more features. Therefore, the reading function dataset needed to be modified in the data input stage. For the selected dataset folder, the corresponding front and back image data were extracted, processed, and enhanced. The torch.cat function was then used to concatenate them by channel, combining two three-channel tensors into one six-channel tensor. This process was carried out by the modified data reading function, mydataset.

(2) Improvement of the convolutional layer

The convolutional layer was improved by merging RGB three-channel images of front and back sides of tobacco leaves into a six-channel image. This six-channel image was then processed by n six-channel convolutional kernels

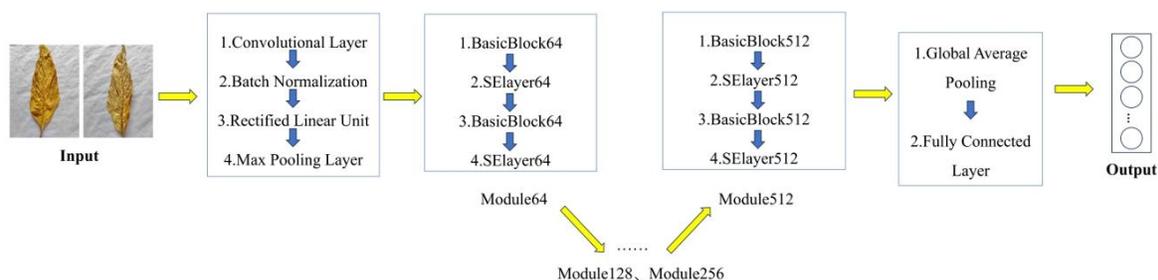


Figure 1. Evolution ResNet (Evo-ResNet) algorithm process.

to generate n feature maps, which were finally combined into an n -channel feature map output. To accommodate the six-channel convolutional layer, the final fully connected layer was adjusted and set its output channels to n .

(3) SE attention mechanism

Given the richer information content in the six-channel input compared to the three-channel input, this research integrated the SE attention mechanism after each BasicBlock to help the better network's utilization of information. With the six-channel input containing the information of both front and back tobacco leaves images, SE attention mechanism enabled the network to effectively comprehend these data by autonomously learning crucial information and enhancing the feature representation of channels. As a result, the network de-emphasized less relevant features, bolstering its generalization prowess. Additionally, to alleviate overfitting, the original ResNet architecture was refined by decreasing the number of BasicBlocks per layer to 2, optimizing parameter usage. The Evo-ResNet model utilized mydataset function to read the image data, and the six-channel tensor was input into modified convolutional layers. Furthermore, SE layers were added after each BasicBlock module, and the final fully connected layer was adjusted to output the corresponding number of categories.

Model establishment and training

The programming software used in this study included PyCharm (JetBrains, Prague, Czech Republic), Jupyter Notebook, an open-source tool primarily managed by the Jupyter

community, and Origin 2024 (OriginLab Corporation, Northampton, MA, USA). The hardware configuration and specifications included Windows 10 System, 64 GB RAM, Intel Core™ i7-10700K CPU @ 3.80GHz, NVIDIA GeForce RTX 3080 with 10 GB memory GPU. This study leveraged the Pytorch framework to establish a custom data loader, mydataset, and configured a six-channel ResNet model. The cross-entropy loss function and Adam optimizer with a learning rate α of 0.00001 were used in this study, while the other parameters were kept at their default values. The model underwent 100 epochs of training and validation cycles followed by final testing.

Model evaluation

The loss-value was determined using cross-entropy, which was a concept in information entropy theory that calculated the difference between the predictions and the true labels for all categories as below.

$$\text{Loss - value} = -\sum_{i=1}^C y_i \log(p_i) \quad (1)$$

where C was the total number of categories in the sample. y_i was the true labels of samples with the label for the true category as 1 and the rest as 0. p_i was the model's predicted probability that a sample belonged to category i . The accuracy indicated the ratio of accurately classified instances to the total, serving as the most prevalent evaluation criterion for classification tasks with the calculation formulated as follows.

$$\text{Accuracy} = \frac{\text{TP} + \text{TN}}{\text{TP} + \text{TN} + \text{FP} + \text{FN}} \quad (2)$$

where T was a correct prediction (True). F was an incorrect prediction (False). P was the predicted results that were positive (Positive). N was the predicted results that were negative (Negative). The recall was the probability of a true positive sample that was correctly predicted as positive.

$$\text{Recall} = \frac{\text{TP}}{\text{TP} + \text{FN}} \quad (3)$$

The precision was the probability of true positive samples among samples predicted as positive.

$$\text{Precision} = \frac{\text{TP}}{\text{TP} + \text{FP}} \quad (4)$$

F1-score was the harmonic mean of recall and precision, which made a balancing judgment between model accuracy and completeness and was shown below.

$$\text{F1-score} = 2 \frac{\text{Precision} \times \text{Recall}}{\text{Precision} + \text{Recall}} \quad (4)$$

Recall, precision, and F1-score were evaluation metrics proposed for binary classification, and each held unique values in such problems, while in multiclassification scenarios, each classification corresponded to its value. These metrics were averaged with a macro method in this study, specifically for precision shown below.

$$P_{\text{macro}} = \frac{1}{n} \sum_{i=1}^n P_i \quad (5)$$

where P_i was the precision of i -th class. N was the total number of classifications. Higher values of the aforementioned four evaluation metrics indicated superior performance.

Results

Training results of Cuibi 1 and Yunyan 87

The two varieties exhibited a similar trend on the training set with smooth progression beyond 80 epochs (Figure 2A). Likewise, a comparable trend between the two varieties on the validation set was demonstrated, showing convergence after 30 epochs with the accuracy of Cuibi 1 slightly exceeding that of Yunyan 87 (Figure 2B). This difference was attributed to the noisier nature of the flue-cured tobacco data for Yunyan 87 compared to Cuibi 1. These findings suggested that the Evo-ResNet model demonstrated greater stability.

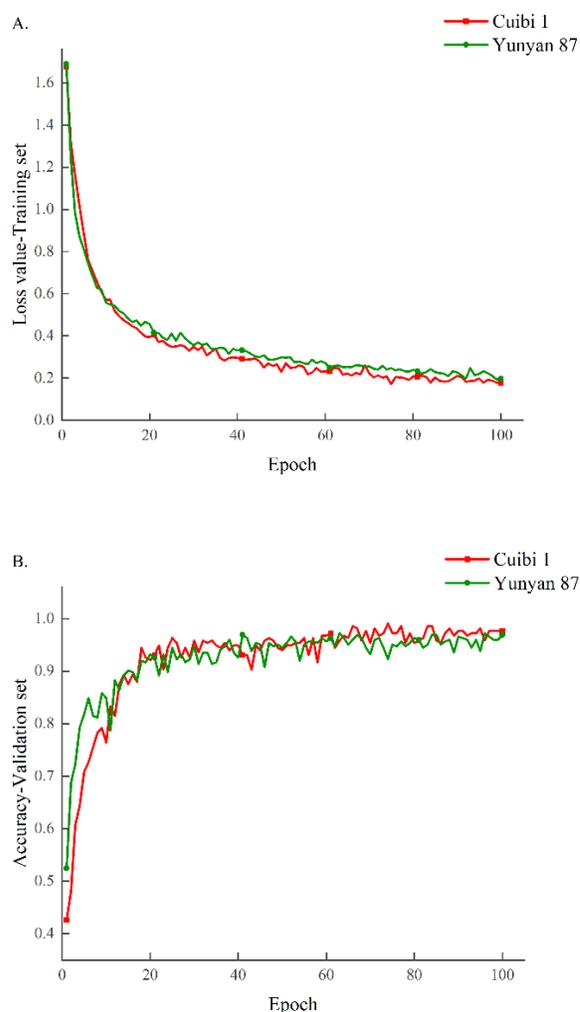


Figure 2. Changes in the loss value of the training set (A) and the accuracy of the validation set (B) of the flue-cured tobacco leaf images.

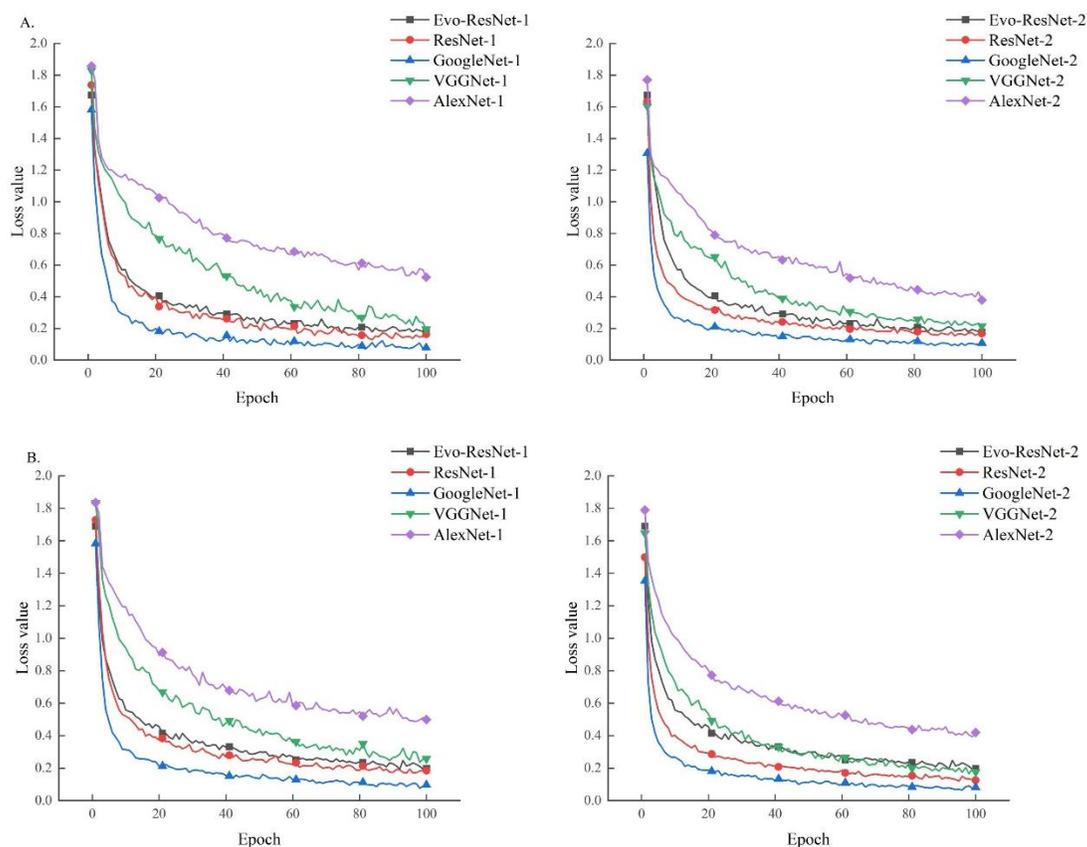


Figure 3. The loss values of training set of Cuibi 1 (A) and Yunyan 87 (B) flue-cured tobacco leaf images using five different network models.

Four evaluation metrics had high values on the test set for the two varieties of flue-cured tobacco leaves with accuracy of 95.83% and 99.07%, recall of 95.83% and 99.07%, precision of 96.16% and 99.10%, and F1-score of 95.78% and 99.08% for Cuibi 1 and Yunyan 87, respectively. These results underscored the model's strong performance in grade classification.

Comparative evaluation of multiple methods

To assess the efficacy of the proposed model, the simultaneous training with conventional three-channel network models including ResNet, GoogLeNet, VGGNet, and AlexNet was conducted in two different ways with only front images of flue-cured tobacco leaves for training, validation, and testing as experiment-1, and both front and back images as experiment-2. The models involved in experiment-1 were denoted as ResNet-1, GoogLeNet-1, VGGNet-1, and AlexNet-1, while the models involved in

experiment-2 were denoted as ResNet-2, GoogLeNet-2, VGGNet-2, and AlexNet-2. The results showed that the loss values of Cuibi 1 and Yunyan 87 network models in 100 epochs exhibited similar loss patterns in both ResNet and Evo-ResNet architectures in both experiments with the stabilization occurring approximately after epoch 80. Notably, GoogLeNet achieved the lowest loss than that of other models, while AlexNet had the highest loss than that of other models after 100 epochs (Figure 3). The analysis results of different network models in two experiments demonstrated that, for Cuibi 1, the loss values were similar between ResNet-1 and ResNet-2, as well as between GoogLeNet-1 and GoogLeNet-2, while the loss values of VGGNet-2 and AlexNet-2 were lower than that of VGGNet-1 and AlexNet-1, respectively. For the Yunyan 87, training loss values of all four networks in experiment-2 were lower than those in experiment-1.

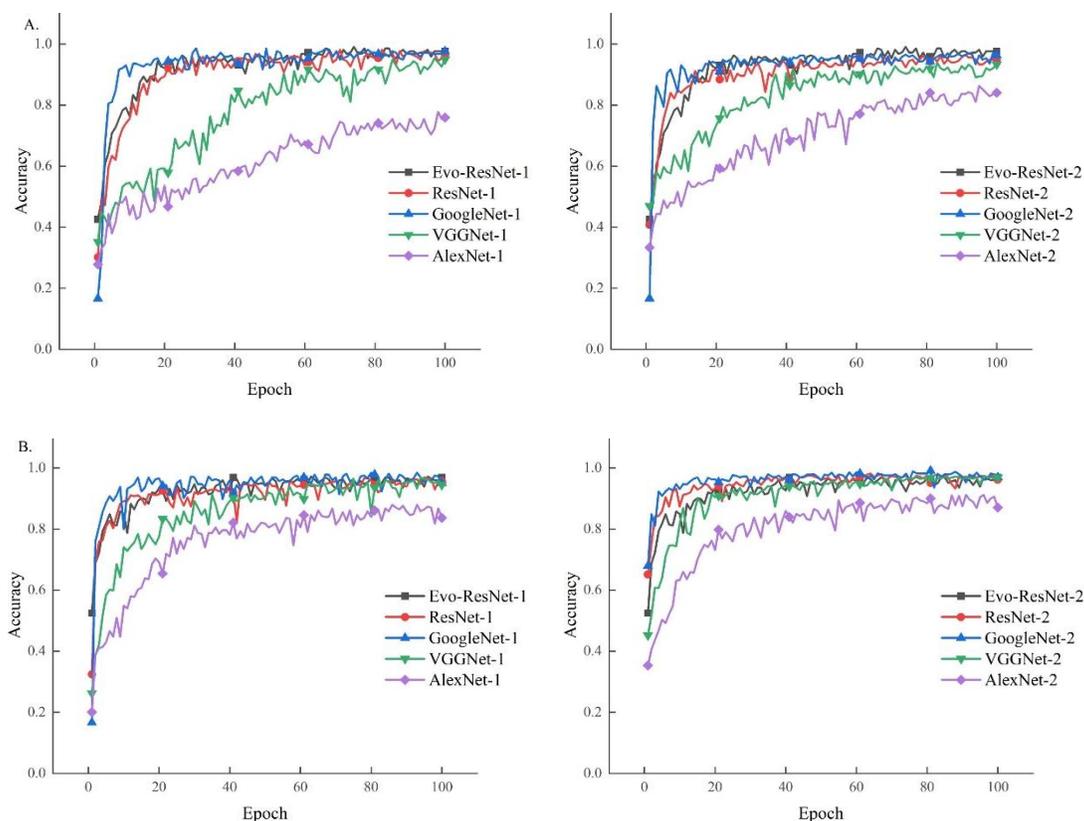


Figure 4. The accuracy value of validation set of Cuibi 1 (A) and Yunyan 87 (B) flue-cured tobacco leaf images using five different network models.

Analysis results of validation set

The changes in accuracy of each network model over 100 epochs on the validation sets of Cuibi 1 and Yunyan 87 were recorded (Figure 4). It was evident that there was similar accuracy for ResNet, GoogleNet, and Evo-ResNet regardless of model training using experiment-1 or experiment-2. However, accuracy metrics of VGGNet and AlexNet were relatively lower than the others.

Analysis results of test set

The evaluation metrics for Cuibi 1 and Yunyan 87 using Evo-ResNet and four conventional CNN models were shown in Table 3. Evo-ResNet demonstrated superior performance compared to the other four conventional CNN models, especially for Cuibi 1 variety. However, its performance was slightly lower than GoogleNet and ResNet for Yunyan 87. Three-channel models in experiment-2 resulted in worse evaluation

metrics compared to that in experiment-1 for Cuibi 1, which might be attributed to the increased data of front and back images that introduced more noise during feature extraction. Therefore, adding SE attention mechanism could enhance the network's ability to reduce learning irrelevant feature, improving the generalization ability and prediction effect. For Yunyan 87, all classification evaluation metrics of three-channel models in experiment-2 were better than those in experiment-1.

The discriminative ability of the Evo-ResNet network model

To evaluate the discriminative capability of the proposed Evo-ResNet model, the classification results on the test set among Evo-ResNet, ResNet, GoogleNet, VGGNet, and AlexNet were compared. To illustrate the discriminative capability of algorithm, the error rate was used to evaluation the recognition ability of algorithm,

Table 3. The classification evaluation of test set for flue-cured tobacco leaf images using five different network models.

Image	Model	Cuibi 1				Yunyan 87			
		Accuracy	Recall	Precision	F1-score	Accuracy	Recall	Precision	F1-score
Front and back	Evo-ResNet	95.83%	95.83%	96.16%	95.78%	99.07%	99.07%	99.10%	99.08%
Front	ResNet	93.98%	93.98%	94.79%	93.92%	98.77%	98.77%	98.79%	98.76%
	GoogLeNet	94.44%	94.44%	95.18%	94.42%	98.15%	98.15%	98.19%	98.15%
	VGGNet	94.91%	94.91%	95.79%	94.96%	97.22%	97.22%	97.46%	97.22%
	AlexNet	78.70%	78.70%	81.18%	78.26%	92.28%	92.28%	92.58%	92.20%
Front and back	ResNet	92.82%	92.82%	93.47%	92.78%	99.23%	99.23%	99.24%	99.23%
	GoogLeNet	92.36%	92.36%	92.93%	92.17%	99.69%	99.69%	99.69%	99.69%
	VGGNet	91.67%	91.67%	92.41%	91.60%	97.84%	97.84%	97.96%	97.83%
	AlexNet	85.19%	85.19%	86.65%	84.88%	94.60%	94.60%	95.14%	94.58%

which was the number of misclassified divided by the total number for each grade. The lower the error rate, the stronger the recognition ability. The results showed that, in experiment-1, the numbers of misclassified images in the classification process of ResNet, GoogleNet, and VGGNet were 13, 12, 11 for Cuibi 1 with an error rate of 6.02%, 5.56%, 5.09%, respectively, while the error rate of AlexNet reached 21.30% with 46 misclassified images. For Yunyan 87, the error rates of three-channel models were lower with the numbers of misclassified images in ResNet, GoogleNet, VGGNet, and AlexNet being 4, 6, 9, 25 and the corresponding error rates of 1.23%, 1.85%, 2.78%, 7.72%, respectively. In experiment-2, the numbers of misclassified images of ResNet, GoogleNet, VGGNet, and AlexNet were 31, 33, 36, and 64 for Cuibi 1 with error rates of 7.18%, 7.64%, 8.33%, and 14.81%, respectively. For Yunyan 87, the numbers of misclassified images of ResNet, GoogleNet, VGGNet, and AlexNet were 5, 2, 14, and 35, respectively, with the corresponding error rates of 0.77%, 0.31%, 2.16%, and 5.40%. The classification error rate of the proposed Evo-ResNet model was 3.24% with 7 misclassified images in the Cuibi 1 dataset. For the Yunyan 87 dataset the error rate decreased to 0.93% with only 3 misclassified images. The results demonstrated that the error rate of Evo-ResNet was just slightly higher than ResNet-2 and GoogLeNet-2 for Yunyan 87. Nonetheless, Evo-ResNet still outperformed the other models.

Generalization ability test of network models

The validation and test set accuracies of the Evo-ResNet and ResNet models for Cuibi 1 and Yunyan 87 were compared to assess the generalization ability of the Evo-ResNet model. The accuracy of the validation set was the highest value in 100 epochs. For Cuibi 1, the accuracy difference between validation and test set of the Evo-ResNet network model was 3.24%, while the ResNet model showed a larger difference of 4.17% in experiment-1 and 3.71% in experiment-2, respectively. For Yunyan 87, the accuracy difference between validation and test set of Evo-ResNet model was 1.85%, while the difference of ResNet model was 1.86% in experiment-1 and 1.85% in experiment-2. The accuracy difference between the validation and test sets of Evo-ResNet network model was the lowest for Cuibi 1 and Yunyan 87 datasets.

Ablation study of the SE module

In the Evo-ResNet model network structure, interdependencies between model channels were found. Therefore, the SE module was added to Evo-ResNet model (Evo-ResNet-SE) for feature recalibration and performance improvement. The results showed that the accuracy, recall, precision, and F1-score of Evo-ResNet-SE model for Cuibi 1 were 95.83%, 95.83%, 96.16%, and 95.78%, respectively. In contrast, the evaluation metrics of Evo-ResNet model were 94.44%, 94.44%, 95.03%, and 94.43% in accuracy, recall, precision, and F1-score, respectively. For Yunyan 87, the Evo-ResNet-SE model had substantial improvements in all evaluation metrics with the accuracy of 99.07%, the recall of 99.07%, the

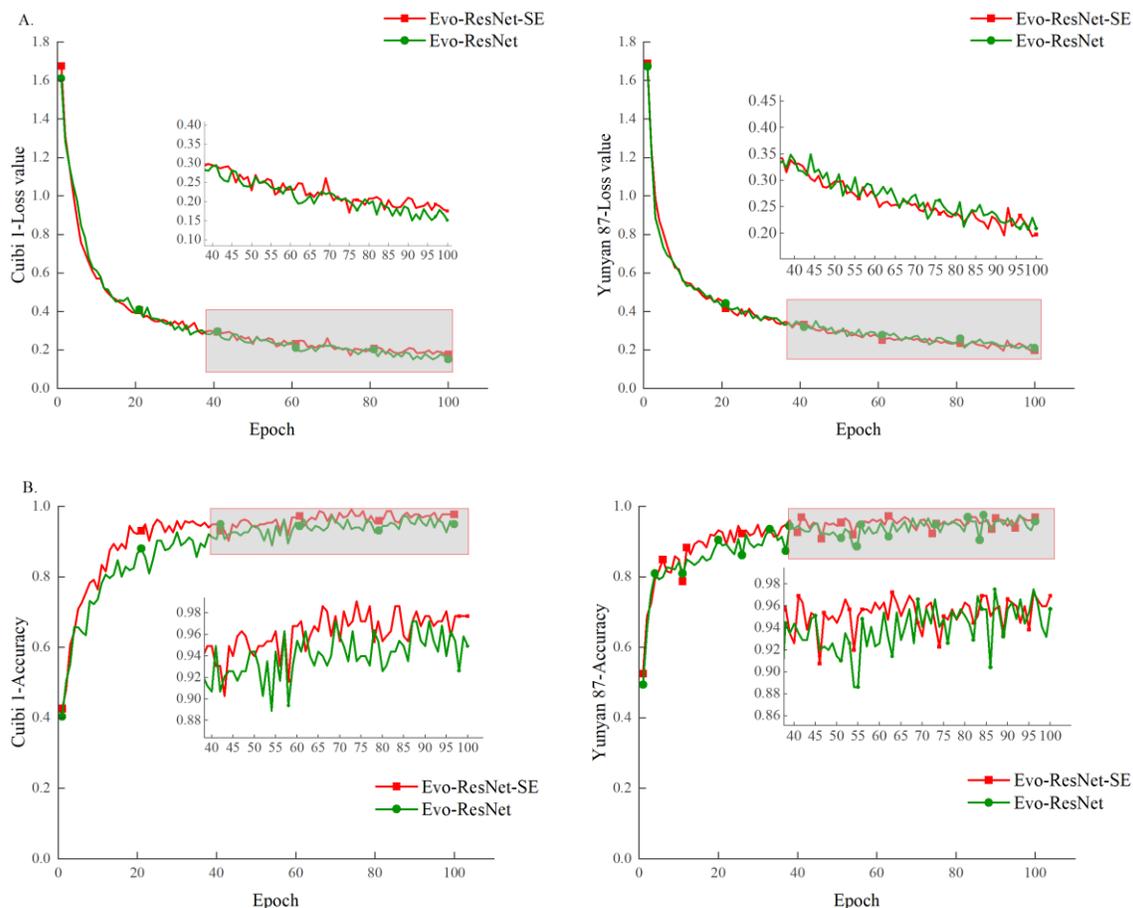


Figure 5. The results of SE module ablation study for Evo-ResNet in training set (A) and validation set (B).

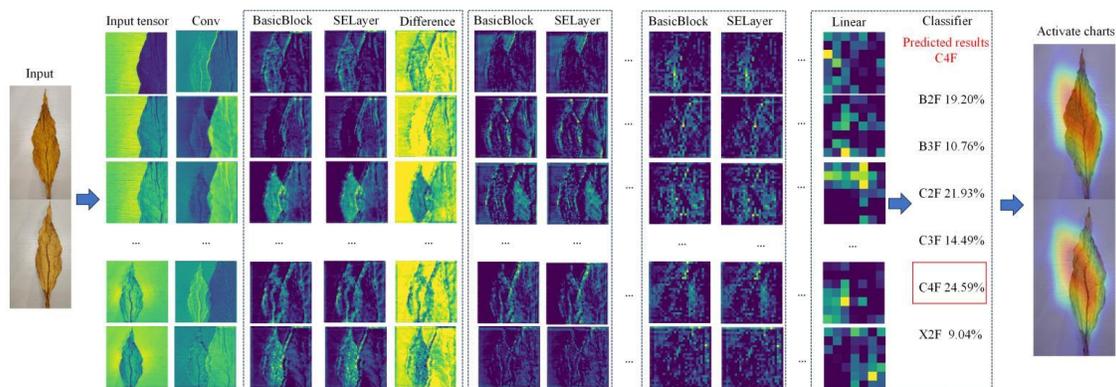


Figure 6. Flowchart of feature visualization for flue-cured tobacco leaf image based on Evo-ResNet model.

precision of 99.10%, and the F1-score of 99.08%, while the metrics of the Evo-ResNet model were 96.30% for accuracy, recall, and F1-score, and 96.78% for precision. The network models with

SE module (Evo-ResNet-SE) and without SE module (Evo-ResNet) showed a consistent trend on the training set, both showing a lower loss value in the Yunyan 87 dataset (Figure 5A). The

accuracy in validation set demonstrated greater fluctuations with the Evo-ResNet-SE model achieving higher accuracy compared to the Evo-ResNet model (Figure 5B). The ablation analysis of the SE module revealed its key role in improving the performance of Evo-ResNet network model.

Image recognition effect based on Evo-ResNet model

In the practical application of grade identification for flue-cured tobacco leaves, the visual variations of image features across layers in the Evo-ResNet model were shown in Figure 6. The feature heatmaps of each network layer showed that the shallow feature maps suggested that the network had extracted characteristics from both front and back images of flue-cured tobacco leaves. The final activation map showed that the model's attention to leaf features was focused on the main vein region, the darker colored wrinkled region, and the tobacco texture, which indicated that the proposed model had effectively utilized both the front and back features of the flue-cured tobacco leaves, enhancing its performance in grade classification.

Discussion

Flue-cured tobacco classification significantly influences the tobacco industry's development. Manual grading is time-consuming and labor-intensive with unstable accuracy rates, especially during the acquisition of flue-cured tobacco leaves. Artificial and intelligent grading mostly relies on the front side of flue-cured tobacco. Recently, there has been growing attention on the back side of flue-cured tobacco. Zhang *et al.* studied color differences between front and back sides of tobacco leaves sourced from Tongren, Zhumadian, Henan, China and Ji'an and assessed their impact on sensory evaluation quality [18]. Li *et al.* investigated color discrepancies between the front and back sides of tobacco leaves sourced from different production areas of the Yunnan's specialty flue-cured tobacco variety 'Honghua Dajinyuan' [19]. Lu *et al.* initially

introduced the classification model to identify front side images by front and back flue-cured tobacco leaf, and then graded these images [12]. This study processed both front and back sides images of flue-cured tobacco leaves concurrently to improve the accuracy of intelligent grading. During intelligent grading of flue-cured tobacco leaves, image collection has usually been achieved by transporting leaves to designated stations for automated capture by HD or hyperspectral cameras. Nonetheless, these equipment-intensive approaches are expensive, and conveyor belt speed fluctuations during assembly line imaging may introduce blurred views. Moreover, external factors like leaf angle, position, and lighting variations also affect image quality consistency. The widespread use of high-resolution mobile phones has facilitated image capturing and precise arrangement of flue-cured tobacco, thereby increasing interest in computer vision research using mobile phone images. Some researchers employed an oven-drying method to measure leaf moisture content using image data collected *via* cell phones. The grayscale histogram image processing techniques were utilized to extract leaf color eigenvalues and analyze their correlation with moisture content [20]. Cell phone-based image collection has broader applications in the intelligent recognition of fresh tobacco maturity. Wang *et al.* captured images of five maturity levels (M1 to M5) of Cuibi 1 (CB-1) fresh tobacco leaves from the upper, middle, and lower sections of the leaves using cell phones and trained these data with a lightweight YOLO network to establish a maturity recognition model for five maturity levels of tobacco leaves [21]. In this study, the application of cell phone-captured images in intelligent grading of flue-cured tobacco was explored and a convolutional neural network model was constructed. The approach of using cell phones for data collection demonstrated its potential for future applications in production processes. This method took advantage of the ubiquity and versatility of modern phone cameras and avoided the need to adhere strictly to specific cell phone models or pixel requirements. With further development, it

could evolve into a dedicated mobile application for flue-cured tobacco procurement with significant practical value. This study also improved the intelligent classification algorithm for flue-cured tobacco based on ResNet. Comparative analysis with traditional algorithms such as GoogLeNet, VGGNet, and AlexNet showed that GoogLeNet outperformed Evo-ResNet, while AlexNet was the least effective in grading flue-cured tobacco images. The proposed method held promise for further optimization of GoogLeNet, VGGNet, and AlexNet. In the future, the following two aspects may be focused, which include adjusting the model structure to reduce the parameter count when using both front and back images of flue-cured tobacco leaves simultaneously that increases the dataset volume and potentially affects training speed and improving model accuracy by enhancing the quality of training dataset because expert dataset obtained from manual grading may contain mislabeled samples. Despite attempts to identify these by discriminative analysis methods, the results are still unsatisfactory. Image processing methods could extract features of flue-cured tobacco and detect mislabeled samples. Based on comprehensive prior analysis, there is limited intelligent grading research that relies solely on single cell phone-captured flue-cured tobacco leaf images and integrates both front and back image data to improve classification performance. This research introduced an enhanced algorithm that utilized a six-channel ResNet convolutional neural network as input. In the Evo-ResNet model, corresponding datasets were created to ensure the one-to-one correspondence of front and back image data inputs. A program was devised to load front and back image data separately and concatenate them into a six-channel tensor along the channels. Subsequently, the input channels of the initial convolutional layer in ResNet were set to six to accept a six-channel tensor input. Additionally, an SE attention mechanism was integrated, and the output channel number of the fully connected layer was set to six. The generalization capability of the Evo-ResNet was outstanding, enabling more precise intelligent

grading of flue-cured tobacco, providing a valuable reference for the rapid and accurate classification during flue-cured tobacco purchasing.

Acknowledgements

This research was funded by the Science and Technology Projects of Fujian Province Company of China National Tobacco Corporation (Grant No. 2022350000240064).

References

1. Chen FL, Sun HQ, Zheng SQ, Guo L, Mu DS. 2012. Streamlining grading efficiency of flue-cured tobacco leaves. *Journal of Mountain Agriculture and Biology*. 31(03):228-230.
2. Zhuang ZZ, Zhu SP, Sun XJ, Zuo YQ. 2016. On method of tobacco automatic grouping based on machine vision. *Artif Intell Sci Eng*. 41(04):122-129.
3. Cho HK, Paek KH. 1997. Feasibility in grading the burley type dried tobacco leaf using computer vision. *J Biosyst Eng*. 22(1):30-40.
4. Li SJ, Pan X, Chen XZ, Zhu JY, Wu BZ, Xie XF, *et al.* 2021. Comparison of tobacco grading methods based on hyperspectral information. *Yancao Keji*. 54(10):82-91.
5. Zhang H, Zhang WW, Zhang YY, Li SJ, Shen JY, Zhu JY, *et al.* 2022. Research on classification method of flue-cured tobacco based on fusion of hyperspectral and texture features. *Zhongguo Yan Cao Xue Bao*. 28(03):72-80.
6. Zhang F, Zhang XH. 2011. Classification and quality evaluation of tobacco leaves based on image processing and fuzzy comprehensive evaluation. *Sensors*. 11(03):2369-2384.
7. Zhang Y. 2011. A review on nondestructive testing technique of flue-cured tobacco based on purchase quality. *Hubei Agric Sci*. 50(07):1297-1300.
8. Mao PJ, Liu L, Wang J, Hu CY. 2010. A study on gray relational analysis of many factor weights in tobacco leaves classification. *Adv Mat Res*. 139-141:1728-1731.
9. Yao XL, He FQ, Ping A, Luo H, Guan QM. 2018. Leaf tobacco grading method based on PCA-GA-SVM. *Yancao Keji*. 51(12):98-105.
10. Liang CY, Wang CC, Ding HN, Zhang JY, Huang YZ, Wang JP. 2022. Research on rapid grading system of tobacco leaf based on machine vision. *J Henan Sci Technol*. 41(11):7-12.
11. Wang SX, Yun LJ, Ye ZX, Wang YB. 2020. A tobacco leaf grading processing algorithm based on convolutional neural network. *J Yunnan Univ Nat Sci Ed*. 29(01):65-69.
12. Lu MY, Zhou Q, Jiang SW, Wang C, Chen D, Chen TE. 2022. Flue-cured tobacco leaf grading method based on deep learning and multi-scale feature fusion. *J Chinese Agric Mech*. 43(01):158-166.

13. Pan M, Wang ZY, Li GH. 2022. Study of construction and application of grading model of tobacco leaf in tobacco area of south Anhui based on depth learning. *Mod Agric Sci Technol*. 2022(14):167-170.
14. Hou Q, Zhou D, Feng J. 2021. Coordinate attention for efficient mobile network design. *IEEE/CVF CVPR*. 2021:13708-13717.
15. Chen SA, Zhao HR, Shentu HQ, Deng YX, Yang J, Zhou D, *et al*. 2023. Research on grade determination of flue-cured tobacco leaves based on residual network. *S Agric Machi*. 54(01):179-183.
16. State Tobacco Monopoly Administration: Flue-cured tobacco: GB2635-1992. Beijing: China Standards Press. 1992:3-5.
17. Huang BR, Fan ZF, Wang F, Jiang YX, Ma XG, Xiao GL, *et al*. 2024. Intelligent grading of flue-cured tobacco based on VGG16-DenseNet integrated model. *Chin Tob Sci*. 45(3):102-112.
18. Zhang YA, Ma CJ, Liu LB, Zhang MQ, Bao KX. 2020. Color of both sides of flue-cured tobacco leaves from different areas and its effect on sensory quality. *J Anhui Agric Sci*. 48(13):198-201.
19. Li YR, Li FL, Yang WK, Liao ZL, Fan JH, He F, *et al*. 2023. Analysis on chromaticity value characteristics of Hong Hua Dajin Yuan flue-cured tobacco leaves. *J Shanxi Agric Sci*. 51(8):888-895.
20. Zhang JL, Song ZY, Han WT, Shang XM, Liu YK. 2019. Nondestructive testing of tobacco leaf water status by using digital RGB images. *J Chinese Agric Mech*. 40(5):62-68.
21. Wang RQ, Zhang BH, Gu G, Shen SJ, Lin XL, Lin JF, *et al*. 2023. Recognition model of tobacco fresh leaf maturity based on YOLOv5. *Zhongguo Yan Cao Xue Bao*. 29(2):46-55.