International Journal of Innovative Computing, Information and Control
Volume **22**, Number **3**, June **2026**

pp. **783–799**

TOBACCO LEAF GRADING METHOD INTEGRATING MULTI-ATTENTION FINE-GRAINED ANALYSIS WITH COLOR AND TEXTURE FEATURES

Mingdeng Fan[1], Dayong Xu[2], Huajie Li[3], Xinfeng Li[1], Guojin Xu[1]
Dengyan Luo[3], Danqiang Fu[1], Chaochao Wang[3], Chen Chen[4]
and Honghai Wu[1,*]

[1]Fujian Longyan Golden Leaf Double Roasting Co., Ltd.
No. 305-1, Yongding District, Longyan 364100, P. R. China
{ 528870923; 1848139573 }@qq.com; { quhn1234; w52199987 }@163.com
*Corresponding author: 3554278179@qq.com

[2]Key Laboratory of Tobacco Processing
Zhengzhou Tobacco Research Institute of CNTC
No. 2, Fengyang Street, Zhengzhou 450000, P. R. China
1282927884@qq.com

[3]Technical Center of Fujian China Tobacco Industry Corporation
No. 118, Lianyue Road, Siming District, Xiamen 361021, P. R. China
{ 2080775578; 1919567099; 396200648 }@qq.com

[4]College of Electrical and Information Engineering
Zhengzhou University of Light Industry
No. 5, Dongfeng Road, Zhengzhou 450000, P. R. China
cc875545317@163.com

Received July 2025; revised November 2025

Abstract. *At present, the current national standards for China's tobacco industry classify tobacco leaves into 42 grades based on 7 indicators. However, these indicators lack a unified quantitative standard, resulting in prominent problems of strong subjectivity and low efficiency in grading. To solve the issues of difficult image feature extraction caused by the curling and folding of redried tobacco leaves, and limited classification accuracy restricted by small inter-class differences during the sorting and grading process of redried tobacco leaves, this article proposes an intelligent tobacco leaf grading method that integrates multi-attention fine-grained features with color and texture features. Firstly, the RGB/HSI dual-space color statistical strategy, Local Binary Pattern (LBP), and Gray-Level Co-occurrence Matrix (GLCM) are used to extract color and texture information, respectively. Meanwhile, a fine-grained classification network model based on multi-attention sampling is designed, which utilizes the attention mechanism to capture global features and local texture details, making up for the defect that traditional convolutional networks are insufficient in distinguishing subtle features. Finally, the Particle Swarm Optimization (PSO) algorithm is introduced to optimize feature weights, and a multi-model integration framework is constructed by fusing color and texture features. In addition, an image dataset of tobacco leaves at various grades is established. Experimental results show that the fine-grained classification network fusion model designed in this paper achieves recognition accuracies of 84.9% and 79.6% for binary classification and five-class classification, respectively, both of which are superior to those of the ResNet and InceptionV3 models. The intelligent tobacco leaf sorting and grading system developed based on this method realizes fully automatic grading of tobacco leaves and effectively improves production efficiency.*

DOI: 10.24507/ijicic.22.03.783

Keywords: Intelligent tobacco leaf grading, Color and texture information, Multi-attention sampling, Attention mechanism, Multi-model integration framework

1. **Introduction.** Tobacco leaves, as the primary raw material for cigarette production, must undergo multiple processing stages and scientific blending to manufacture cigarettes that meet consumer demands. Grading enables the differentiation of tobacco leaves based on quality, allowing the cigarette industry to tailor blending formulations according to the characteristics of each grade, thereby ensuring consistent and stable product quality [5]. To ensure consistent cigarette product quality, rigorous control of tobacco leaf quality is imperative. Consequently, establishing scientific, rational and standardized tobacco leaf grading criteria are critically important for cigarette manufacturing.

In the quality control system of the global tobacco industry, the United States, as a pioneer in the standardization of tobacco leaf grading, has incorporated 12 indicators into its grading standards. It also combines these standards with a production area database to realize the automatic association between grading results and tobacco leaf purchase prices [1]. However, its database only covers local production areas and cannot be adapted to the grading of imported tobacco leaves. Japan has developed a dual verification system that initially categorizes tobacco leaves using machine vision, and then rapidly detects nicotine content via near-infrared spectroscopy to refine grading results, achieving a grading accuracy rate of 90% [2]. However, the system's detection time is relatively long, requiring 30 seconds per leaf, which struggles to meet the demands of large-scale production line efficiency. Germany focuses on the modularization of grading equipment [3]. Its developed detachable visual inspection unit can quickly replace inspection modules according to different tobacco types such as flue-cured tobacco and burley tobacco, adapting to the grading needs of multiple categories. Although German modular equipment is flexible, it requires recalibration after module replacement, which increases operational complexity.

China is the world's largest tobacco producer and consumer, with annual production accounting for approximately 30% of global tobacco output [4]. China's current national tobacco industry standard classifies tobacco leaves into 42 grades [6] based on seven key factors: maturity, leaf structure, body, oil content, color intensity, length and blemishes. However, these parameters lack unified quantitative definitions and standardized measurements. As a result, during commercial procurement and threshing-redrying processes, tobacco grading still primarily relies on field technicians who make determinations based on reference grading samples, technical specifications, and their professional expertise incorporating visual inspection, tactile evaluation and accumulated experience. However, the manual grading approach presents multiple challenges: low efficiency, high labor intensity, elevated costs, strong subjectivity, inconsistent evaluation standards, quality variability, frequent occurrences of mixed immature/low-grade leaves or incorrect stalk position classification, as well as integrity risks in grading [7]. These limitations make it increasingly difficult to meet the practical demands of modern tobacco production.

At present, an increasing number of researchers at home and abroad have begun to explore automated grading and sorting solutions to replace manual operations in the entire process of tobacco leaf feeding, bale opening, loosening and spreading, single-leaf separation, intelligent identification and grading, as well as sorting and recovery [8, 9, 27]. Within this framework, intelligent tobacco leaf recognition and classification has emerged as a pivotal research focus, driving significant technological advancements in this field. Currently, domestic and international research on intelligent tobacco leaf grading can be primarily categorized into three directions:

1) The approaches based on computer vision technologies and traditional machine learning methods [10, 11, 12, 13, 24, 30]. This methodology involves extracting features from tobacco leaf images, then utilizing manually graded labeled datasets to train classifiers such as multilayer perceptrons, clustering algorithms, and decision trees for classification. Bin et al. [14] proposed a PCA-GA-SVM tobacco grading method incorporating Principal Component Analysis (PCA), Genetic Algorithm (GA) and Support Vector Machine (SVM). Compared to standalone SVM and GA-SVM models, this approach demonstrated 24.86% and 35.64% improvements in recognition rate and grading efficiency, respectively. However, this method relies heavily on manual feature engineering, resulting in limited model generalizability. The performance is constrained by the quality of handcrafted features and may not adapt well to variations in real-world tobacco leaf samples. He et al. proposed an end-to-end Cross-Modal Enhancement Network (CMENet) for the automatic grading of tobacco leaves [15]. In addition to the common reflection images, this network also adopts transmission images to incorporate the thickness information used in manual grading.

2) Grading methods based on spectroscopy technology combined with machine learning. This approach first acquires spectral data from tobacco leaves using Near-Infrared (NIR) or hyperspectral imaging equipment, and then extracts intrinsic features such as oil content and thickness for grading. Li et al. [16] proposed an automatic tobacco leaf grading method based on NIR spectroscopy combined with an Extreme Learning Machine (ELM) algorithm. The external prediction accuracy for upper, middle, and lower-grade tobacco leaves exceeded 90%. Luo [17] investigated the impact of different preprocessing methods and four classification models – Random Forest (RF), ELM, Gradient Boosting Decision Tree (GBDT), and SVM – on grading accuracy using hyperspectral data from various tobacco grades. While spectroscopy-based methods demonstrate high accuracy, they require relatively expensive hardware. Additionally, the extracted features are susceptible to interference from variations in leaf morphology.

3) Methods based on deep learning convolutional neural networks [19, 20]. The method utilizes manually annotated tobacco leaf images to train a convolutional neural network classification model for tobacco leaf classification. Liu et al. [4] proposed a multi-scale feature fusion-based classification method for tobacco leaf grades using RGB images of flue-cured tobacco and deep learning [21, 23, 26]. The approach employs ResNet50 to extract features from tobacco leaf images and incorporates an attention mechanism and FPN, achieving an accuracy of 80.14% in classifying tobacco leaves into seven grades.

The aforementioned methods each have their own advantages and disadvantages, but they are primarily based on analyzing tobacco leaf images or spectral data collected in experimental settings where the leaves are laid flat and either stationary or moving at low speeds. However, in actual automated grading equipment, tobacco leaves move at high speeds of nearly 3 m/s, with various irregularities such as wrinkles, curling, and overlapping. Under these conditions, single-feature or single-model classification methods exhibit low accuracy and fail to meet real-world production requirements [18, 25, 28, 29].

To address these challenges, this paper proposes a multi-attention fine-grained feature and color-texture fusion method for raw tobacco leaf grading. Additionally, an intelligent tobacco leaf grading system is designed based on this method, enabling automatic and precise grading of tobacco leaves. The innovations of this paper are as follows.

1) A multi-angle combined line-scan lighting system was designed to capture images of raw tobacco leaves moving at high speed in curled and wrinkled states, ensuring uniform brightness and shadow-free imaging. Additionally, a graded tobacco leaf image dataset was constructed, incorporating high-quality image information for subsequent analysis and processing.

2) A cascaded intelligent tobacco leaf classification algorithm was designed. This algorithm extracts color and texture information through RGB/HSI dual-space color statistics, Local Binary Patterns (LBP), and Gray-Level Co-occurrence Matrix (GLCM), while leveraging a fine-grained classification network model to capture global morphological features and local texture details. Additionally, a multi-swarm particle optimization algorithm was employed to optimize feature weight allocation, enabling effective fusion of color and texture features to construct the classification model and further improve classification accuracy.

3) Based on the designed classification model, an intelligent tobacco leaf grading system was developed, achieving accurate recognition and automated grading of tobacco leaves across different quality grades.

The remainder of this article is organized as follows. In Section 2, the proposed method is introduced. The establishment of the datasets is described in Section 3. And then, the experimental results and analysis are given in Section 4. Finally, we draw the conclusions in Section 5.

2. The Proposed Method.

2.1. **Overview.** This paper proposes an intelligent tobacco leaf grading model that integrates multi-attention fine-grained features with color-texture characteristics. The model employs the following technical approach. First, RGB/HSI dual-space color statistical features (mean, variance) are used to model the overall color distribution of tobacco leaves. An LBP-GLCM texture feature extraction method is implemented to obtain direction-consistent texture features from wrinkled regions. Second, we design a Multi-scale Attention Sub-Network (MASN) based on SE-Net's channel attention mechanism. This architecture dynamically adjusts feature channel weights while employing a spatial adaptive sampling strategy to focus on critical areas such as vein density. Finally, a particle swarm optimization algorithm adaptively fuses the decision weights of color, texture, and deep features to construct a multi-level classification model specifically designed for deformed tobacco leaves. The model architecture is illustrated in Figure 1.

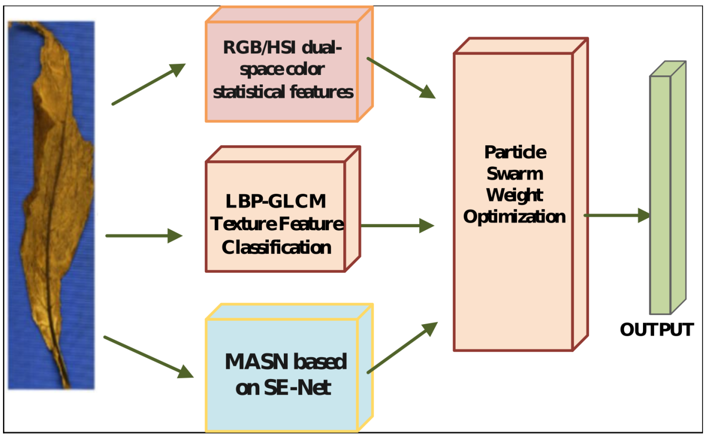


FIGURE 1. The proposed model framework architecture

2.2. **Color characteristics classification.** The color characteristics of tobacco leaves are also an important basis for grouping tobacco leaves, which are related to the identity, oil content, and leaf structure of tobacco leaves. The growth parts of tobacco leaves in different groups are different, so their environmental factors such as light exposure time, ventilation rate, and degree of photosynthesis are not the same, resulting in different colors of tobacco leaves. This article takes the first, second, and third moments of the RGB color space and the channel mean and variance of the HSI color space as color features, and then trains the classifier through SVM to obtain the color feature classification model M_c.

2.3. **LBP-GLCM texture feature classification.** The texture of the tobacco leaf can reflect the leaf structure and the degree of fold flatness to a certain extent, so it can be used as a feature parameter for judging the grading of the tobacco leaf. In this paper, LBP-GLCM is used for the extraction of tobacco leaf texture features. Local Binary Pattern (LBP) generates a binary code describing the local texture pattern by comparing the difference in grey values between each pixel point in an image and its 8 pixels in a 3×3 neighbourhood, which is robust to illumination variations and local deformations, and is suitable for stable texture feature extraction in folded tobacco. Grey Level Co-occurrence Matrix (GLCM): under the pixel spacing of $d = 1$, the co-occurrence probability of grey level values is counted along the four directions of 0°, 45°, 90°, and 135°, and the statistics such as contrast and energy are calculated to quantify the directionality and spatial distribution law of the tobacco texture and to enhance the ability of texture characterization of folded regions. The grey scale covariance matrix itself cannot be used to describe the texture features, it is necessary to use mathematical computation methods to get some parameters from the matrix that reflect the condition of the matrix, in this paper, we extracted four texture features of the tobacco leaf: texture contrast, texture energy, texture entropy, and texture correlation, and the expressions are as follows.

Texture Contrast: It refers to the weighted sum of the squared differences in gray levels and the corresponding co-occurrence probabilities, and serves as an indicator for describing the magnitude of gray level differences between adjacent pixels in tobacco leaf images.

$$Con = \sum_{i=0}^{L-1} \sum_{j=0}^{L-1} (i-j)^2 p(i,j) \tag{1}$$

Texture Energy: It refers to the sum of the squares of the probabilities of elements in the gray-level co-occurrence matrix, and is used to describe the uniformity and regularity of texture in tobacco leaf images.

$$ASM = \sum_{i=0}^{L-1} \sum_{j=0}^{L-1} p(i,j)^2 \tag{2}$$

Texture Entropy: It represents the weighted sum of negative probabilities and the logarithm of probabilities, and is an indicator used to describe the complexity and randomness of tobacco leaf texture.

$$ENT = -\sum_{i=0}^{L-1} \sum_{j=0}^{L-1} p(i,j) \log(p(i,j)) \tag{3}$$

Texture correlation: t represents the result of normalizing the weighted sum of grayscale deviation products and probabilities by the standard deviation, and is used to measure the degree of linear correlation between the grayscale values of adjacent pixels in tobacco leaf texture.

$$IDM = \frac{\sum_{i=0}^{L-1} \sum_{j=0}^{L-1} (i-\mu_i)(j-\mu_j) P(i,j)}{\sigma_i \sigma_j} \tag{4}$$

Here, $P(i, j)$ is an element of the gray-level co-occurrence matrix. i and j are gray levels. L represents the gray level of the tobacco leaf grayscale image, with its value range of

0-255. μ_i and μ_j are the mean values of gray levels i and j, respectively, and σ_i and σ_j are the standard deviations of i and j, respectively.

In the texture feature extraction stage, a two-path fusion strategy is used to improve the robustness of the representation of folded tobacco: firstly, the tobacco image is subjected to LBP processing to generate LBP coded images reflecting the local texture structure; subsequently, multi-directional GLCM analyses are performed on the original grey-scale image and the LBP texture image, respectively, by setting the pixel spacing $d = 1$, and constructing along the four directions of 0°, 45°, 90°, and 135°. Symbiosis Matrix. Statistics such as contrast, energy, entropy and correlation are calculated for each matrix, and the mean and standard deviation of the statistics of the original image and the LBP image in the four directions are fused to construct a 16-dimensional texture feature vector F_t with the ability of spatial-directional joint characterization. Based on the training set of feature vectors, the texture discriminative model is established by the Support Vector Machine (SVM), and the texture direction and spatial distribution characteristics of the creased tobacco leaves are effectively captured.

2.4. **Multi-attention sampling fine-grained classification networks.** Aiming at the challenges of fine-grained classification in which tobacco subclasses have subtle differences and lack significant local structural features, this paper proposes a fine-grained classification network (MASN) based on multi-attention sampling. The method automatically strengthens the key features related to tobacco classification (e.g., vein density, and fold edge orientation) through the channel attention mechanism, combines the spatial adaptive sampling strategy to dynamically focus on the highly discriminative region, and autonomously captures the pixel-level subtle differences without manual annotation, which effectively solves the problems of the traditional method's strong dependence on the tobacco leaf flatness and the insufficient ability of the interclass discriminative ability. As shown in Figure 2, the multi-attention sampling network mainly consists of four parts: convolutional feature map extraction module, attention module, attention map sampling module, and classification module.

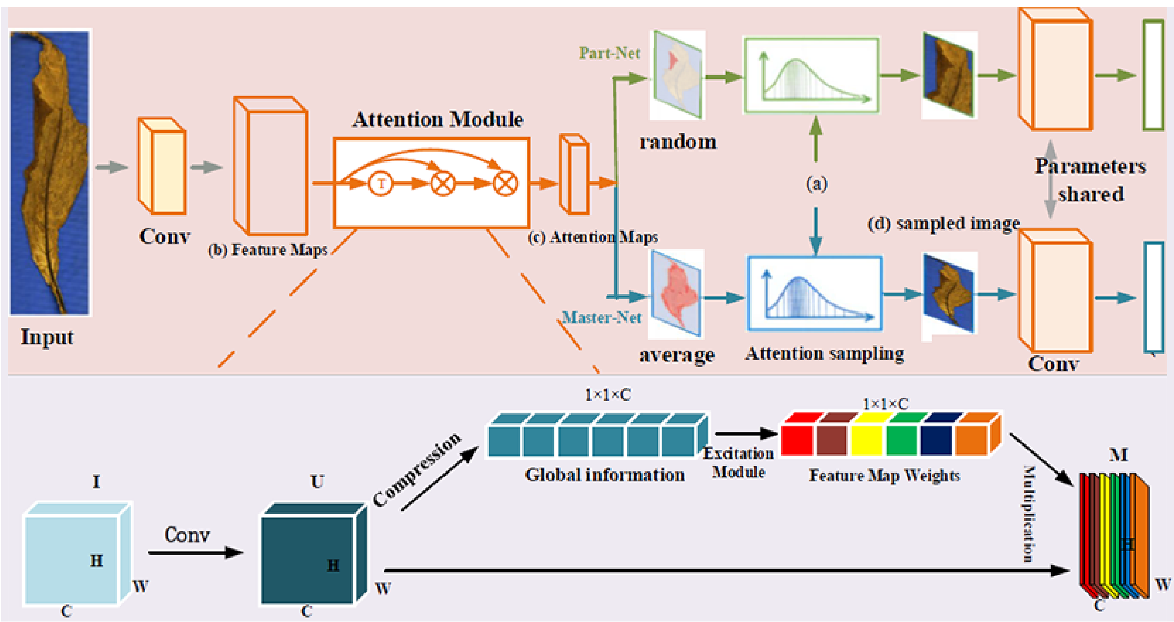


FIGURE 2. Multi-attention sampling fine-grained classification network architecture

Firstly, for the input image I, the depth features are extracted by the convolutional feature map extraction module to obtain the feature map $U \in \mathbb{R}^{H\times W\times C}$ as shown in Figure 2(b), and the feature map U is input into the attention module to be converted into the attention map $M \in \mathbb{R}^{H\times W\times C1}$ as shown in Figure 2(c).

For the processing of the attention graph M, we divide the network into two branches: the part branch (Part-Net) and the structure branch (Master-Net). To learn the fine-grained features of a particular part, we randomly select an attention map M_n $(0 < n < C1)$ and use this attention map to sample the attention of the input image I. This processing flow constitutes the Part-Net branch because it preserves specific details at high resolution. In addition, in order to capture the global structure and include all the important details, we average all the attention maps and do the attention sampling again, and this process constitutes the structure branch (Master-Net). The structure branch (Master-Net) is used to learn the structure preserving global image features and the part branch (Part-Net) is used to learn the fine-grained representation of each detail preserving image.

2.4.1. *SE-Net.* In this article, we adopt the SE-Net channel attention mechanism to strengthen the key cues of tobacco classification through adaptive feature calibration: global compression of input features to generate channel weight vectors, and dynamic enhancement of the response strength of channels with highly discriminative features such as leaf vein texture and folded edges. This mechanism allows the network to autonomously focus on the subtle color difference and texture changes related to grading, and significantly improves the detailed characterization of folded tobacco leaves, and its key structure consists of two parts, namely, compression (Squeeze module) and excitation (Excitation module). The Squeeze module compresses the feature map $U \in \mathbb{R}^{H\times W\times C}$ to $Z_c \in \mathbb{R}^{1\times 1\times C}$, and realizes the conversion of spatial features to global features, and the resulting global features are converted into global features. The obtained global information is used as feature weights, which are operated by global average pooling according to Equation (5):

$$Z_c = \frac{1}{H\times W}\sum_{i=1}^{H}\sum_{j=1}^{W} U_c(i,j) \tag{5}$$

Here, U_c denotes the characteristic map of the C channel.

The Excitation module better fits the complex correlation between channels by amplifying the weights of the most important regions in the feature maps and ignoring the irrelevant parts using a fully connected layer with an activation function, where the vectors represent the degree of importance of each feature map as in Equation (6). The original feature map is motivated by the vector of Equation (7) for continuous iterative updating.

$$s = \sigma(w_2(w_1 z)) \tag{6}$$

$$M_c = S_c \cdot U_c \tag{7}$$

Here, σ denotes ReLU activation function, $w_1 \in \mathbb{R}^{\frac{C}{r}\times C}$, $w_2 \in \mathbb{R}^{C\times\frac{C}{r}}$, and r is the scaling factor.

For Part-Net branching (Part-Net), a channel is randomly selected from the attention graph M and the input image is attentively sampled using this channel as shown in Equation (8).

$$I_d = \mathcal{D}(I, R(M)) \tag{8}$$

2.4.2. *Master-Net.* For structural branching (Master-Net), all the attention graphs are first averaged and then an attention sampling is done, which is calculated as shown in Equation (9) to Equation (11). Firstly, after the average pooling for the attention graph

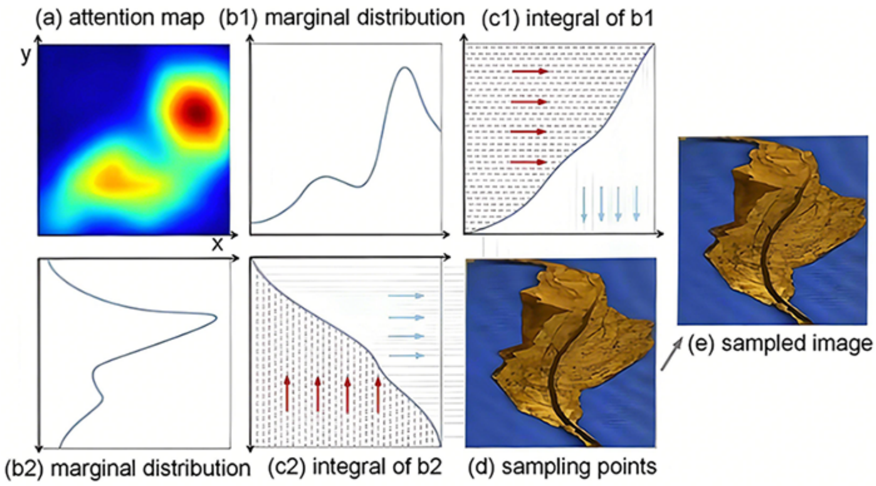


FIGURE 3. Attention sampling module

M is decomposed into two dimensions: x and y, and then the y-axis is fixed first to obtain the maximum value on the x-axis, and then the x-axis is fixed again to obtain the maximum value on the y-axis, and then the integrals are made for these two axes, i.e., Figure 3(c1) and Figure 3(c2) are obtained. After obtaining the integrals, the inverse function (Formula (9)) is applied to obtain the sampling points, as shown in Figure 3(d). Then, the original image is sampled at these points to obtain Figure 3(e).

$$I_s = S(I, A(M)) \tag{9}$$

$$\begin{aligned} f_x(n) &:= \sum_{j=1}^{n} \max_{1 \le i \le W} A(M)_{i,j} \\ f_y(n) &:= \sum_{i=1}^{n} \max_{1 \le j \le H} A(M)_{i,j} \end{aligned} \tag{10}$$

$$S(I, A(M))_{i,j} = I \cdot f_x(i) \cdot f_y(j) \tag{11}$$

where S represents the sampling function, A represents the average pooling operation, and M is the attention graph.

After obtaining the attention sampling module Figure 3(d) and Figure 3(e), these two images are fed to the convolutional neural network ResNet18 to obtain the fully connected outputs, denoted as z_s and z_d, respectively, which are then converted to the probability vectors q_s and q_d by means of the softmax function, with the following softmax formula:

$$q_s^{(i)} = \frac{\exp\left(z_s^{(i)}\right)}{\sum_j \exp\left(z_s^{(j)}\right)} \tag{12}$$

The loss function is calculated using binary cross entropy and soft target cross entropy:

$$L(I_s) = L_{cls}(q_s, y) + \lambda L_{csn}(q_s \cdot q_d) \tag{13}$$

where $L_{cls}(q_s, y)$ is the binary cross entropy, $L_{csn}(q_s, q_d)$ is the soft target cross entropy and λ is the weight parameter. The formula for $L_{cls}(q_s, y)$ and $L_{csn}(q_s, q_d)$ is

$$L_{soft}(q_s, q_d) = -\sum_{i=1}^{N} q_d^{(i)} \log q_s^{(i)} \tag{14}$$

$$L_{cls}(q_s, y) = -\sum_{(i,j)}^{(H,W)} [y \log q_s + (1-y)\log(1-q_s)] \tag{15}$$

2.5. **Particle swarm model optimization strategies.** The particle swarm algorithm is used to optimize the weight values of each model in the following steps.

1) For the training dataset $X = \{x_1, x_2, \ldots, x_n\}$, the initialized weights of the three models of color, texture and multi-attention sampling fine-grained classification network are random and the three model weights satisfy the following update rules.

2) The maximum number of iterations is set to 1000, and the initial speed is set to 0; the individual optimal weights are initialized p_i, with the group optimal weights initialized p_g. The threshold is set to 0.95 for model training.

3) Start the iterative process which includes weights iteration and speed iteration in the following equation:

$$\mathbf{v}_i^{(k+1)} = \omega \mathbf{v}_i^{(k)} + 2\gamma_1 \left(\mathbf{p}_i^{(k)} - \mathbf{w}_i^{(k)}\right) + 2\gamma_2 \left(\mathbf{p}_g^{(k)} - \mathbf{w}_i^{(k)}\right) \tag{16}$$

$$\mathbf{w}_i^{(k+1)} = \mathbf{w}_i^{(k)} + \mathbf{v}_i^{(k+1)} \tag{17}$$

where $i = 1, 2, \ldots, n$ (n is the number of training samples); k is the number of current iterations; w is the inertia weight, and γ_1 and γ_2 are random numbers taking the values $[0, 1]$.

3. **Establishment of Datasets.** A 4K line-array industrial camera was used to acquire images of tobacco samples. The freshly roasted tobacco in its natural state is not flat; the height of different regions of the tobacco is uneven, and they mutually block each other. To solve the problems of uneven image brightness caused by the tobacco moving at high speed while folded and curled, a multi-angle, multi-focus surface combined light source was designed, as shown in Figure 4. By combining light sources with different angles and focusing surfaces, the color of the collected tobacco images becomes stable and the brightness uniform. For 34600 pieces of tobacco samples, online image data acquisition was carried out separately to obtain the image data of each grade of tobacco with labels, and for each grade of tobacco, the image data were repeatedly acquired for five times for model training, validation and testing. The test tobacco samples were raw tobacco grown in Sanming, Fujian Province and purchased in the first baking, which were categorized into six grades, including B01, B03, C01, C03, F03 and X02. The actual grading was performed according to the two models of two choices for B01 and B03, and five choices for B01, C01, C03, F03 and X02.

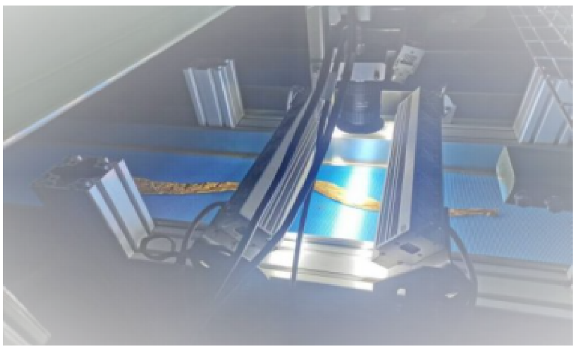


Figure 4. Multi-angle multi-focus surface combined light source

The 34600 tobacco samples used for testing were manually sorted into selected grades by a professional grader, and the grades were divided into a training set T, a validation set V, and a test set S according to 7 : 2 : 1, and the numbers of grades and subsets are shown in Table 1.

TABLE 1. Quantity of each grade of tobacco leaf samples for testing

	Subsets			
Categories	Training set	Validation set	Test set	Total number
B01	4760	1360	680	6800
B03	4305	1230	615	6150
C01	3920	1120	560	5600
C03	4375	1250	625	6250
F03	3486	996	498	4980
X02	3374	964	482	4820
Total				34600

4. Experimental Results and Analysis.

4.1. Evaluation indicators. The original classified correct rate (i.e., tobacco grade pass rate) was used to measure the model accuracy during the model training and testing process, and the classified correct rate is the ratio of the number of correctly classified samples to the number of all samples, as shown in Equation (18).

$$Acc = \frac{TP_1 + TP_2 + TP_3 + \cdots + TP_c}{N} \tag{18}$$

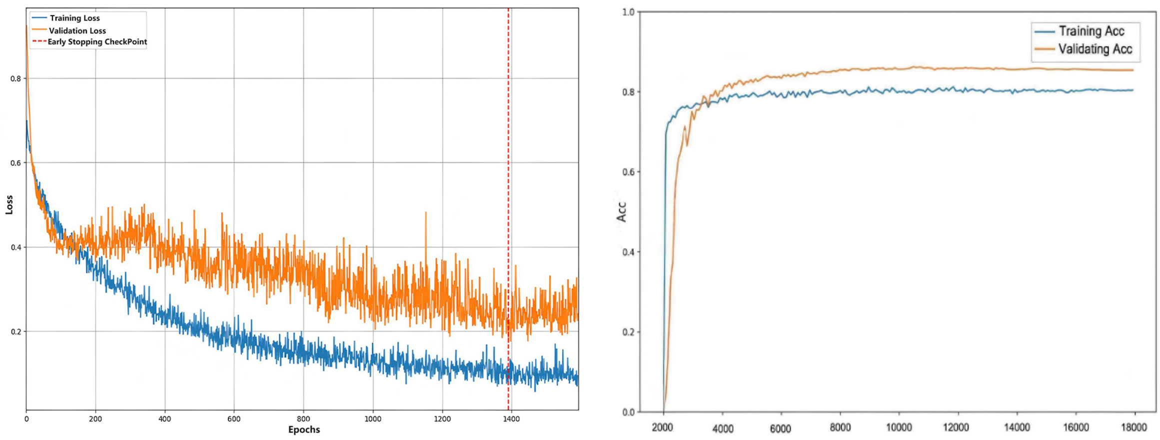
where N is the total number of test samples, c is the number of categories, and TP_c is the number of samples in category c that were correctly categorized.

Meanwhile, the recognition rate index is utilized to judge the model detection performance, and the calculation process is as follows:

$$T = \frac{\text{Correctly identified tobacco leaves}}{\text{Total number of tobacco leaves}} \tag{19}$$

4.2. Model runtime environment. The experimental process is divided into two parts: model training and validation and testing in a real production environment. Among them, the training environment is used to train three single models of color, texture and Multi-Attention Sampling Fine-Grained Classification Network (MASN) model as well as the fusion model. In the training environment, the operating system is Ubuntu 20.04, two A100 graphics cards with 40G RAM each, the deep learning framework is pytorch with version 1.12, and the programming language is python. In the real production environment, the operating system is Ubuntu 20.04, the GPU is GeForce RTX 2080 with 12G RAM, and the programming language is C++. The trained models are deployed for inference optimization using the tensorRT framework. tensorRT is a high-performance GPU inference C++ library developed by NVIDIA, which accelerates the inference of deep learning models by combining layers, optimally selecting convolutional kernels, half-precision operations, and optimizing computation in real time according to hardware.

4.3. Model training. The trained color, texture and Multi-Attention Sampling Fine-Grained Classification Network (MASN) model input image height and width is $800 * 512$, for the training of the Multi-Attention Sampling Fine-Grained Classification Network (MASN) model, the input batch is 48, the training iterations are performed in a way of



(a) Loss training curve (b) ACC training and validation results

FIGURE 5. Attention sampling moudle

fine tuning based on the parameters of the pre-trained model, and a total of 1600 epochs were trained. Stochastic Gradient Descent (SGD) approach was used to optimize the parameters with an initial learning rate of 0.005, and learning rate tuning was carried out by cosine decay. Before training, learning rate warm-up (warm-up) iterations are performed for 2 times. During training, the image data are enhanced with “random cropping” and “random flipping”, and the early stopping mechanism is used to prevent the model from overfitting. Variation of training loss function value with training steps are shown in Figure 5(a). Figure 5(b) shows the results of ACC training and evaluation, and it can be seen that the algorithm in this paper has a high performance in judging the grade passing rate.

4.4. **Model performance analysis.** In order to evaluate the tobacco classification effect of the method proposed in this paper, comparative experiments are conducted on the validation set, which are divided into two groups, B01 and B03 two-classification test and five-classification test, respectively, to compare the recognition classification accuracy of the single color model M_c, the texture model M_t, as well as the general-purpose convolutional neural network models ResNet18, ResNet50, InceptionV3, and the Multi-Attention Sampling Fine-Grained Classification Network (MASN) model, fusion model in the two-classification test and five-classification test, respectively. The results are shown in Tables 2 and 3. It can be seen that the MASN and the particle swarm optimization fusion model proposed in this paper are better than a single feature model and a general convolutional neural network model in terms of recognition accuracy.

TABLE 2. Experimental results of tobacco classification test

Model	Number of test validation set images	Correctly identified quantities	Recognition rate %
Color Model M_c	12950	8223	63.5
Texture Model M_t	12950	8469	65.4
ResNet18	12950	9816	75.8
ResNet50	12950	9997	77.2
InceptionV3	12950	10178	78.6
MASN	12950	10437	80.6
M_c+M_t+MASN	12950	10994	84.9

TABLE 3. Experimental results of tobacco classification test

Model	Number of test validation set images	Correctly identified quantities	Recognition rate %
Color Model M_c	34600	19168	55.4
Texture Model M_t	34600	20863	60.3
ResNet18	34600	24842	71.8
ResNet50	34600	25327	73.2
InceptionV3	34600	25534	73.8
MASN	34600	26780	77.4
M_c+M_t+MASN	34600	27541	79.6

By comparing Tables 2 and 3, it can also be seen that the color and texture model has a significant decrease in the recognition rate at five classifications, this is because for the five classifications of each grade of tobacco in the color and texture characteristics, the differences are much smaller, and a single model has been unable to distinguish between them.

4.5. **Tobacco grading tests.** As described in the previous section, among all the sample tobaccos, they were divided into training set T, validation set V and test set S according to 7 : 2 : 1, in which the test set samples were not used for training. To verify the generalization performance of the model of the method proposed in this paper and the actual testing effect, the tobacco samples of the test set are used to carry out online two-classification and five-classification grading tests on the intelligent sorting and selection grading equipment, and the results of the cyclic testing are shown in Tables 4 and 5. It can be seen that, in the online grading carried out in the actual production environment, the recognition rate of the particle swarm optimization fusion model is close to that of the validation set, which indicates that the algorithm has a good generalization performance and a high level of robustness.

TABLE 4. Experimental results of tobacco classification test

Number	B01 number	Correctly identified quantities	Recognition rate %	B03 number	Correctly identified quantities	Recognition rate %
1	680	565	83.1	615	517	84.1
2	680	551	81.0	615	509	82.7
3	680	561	82.5	615	512	83.2
4	680	569	83.7	615	507	82.4
5	680	559	82.2	615	519	84.4

TABLE 5. Experimental results of tobacco classification test

Number	B01 number	C01 number	C03 number	F03 number	X02 number	Correctly identified quantities	Recognition rate %
1	680	560	625	498	482	2196	77.2
2	680	560	625	498	482	2159	75.9
3	680	560	625	498	482	2167	76.2
4	680	560	625	498	482	2182	76.7
5	680	560	625	498	482	2175	76.4

The recognition rates of the five classified categories are shown in Table 6, which shows that the recognition rates of each category are not balanced, with C01 and C03 having lower accuracy rates, and X02 having a significantly higher accuracy rate than several other categories, which is due to the fact that C01 and C03 themselves are very close to each other, and the difference between the categories is very small.

TABLE 6. Recognition rate results of different tobacco ranks

Sample level	Sample number	Correctly identified quantities	Recognition rate %
B01	3400	2594	76.3
C01	2800	2069	73.9
C03	3125	2262	72.4
F03	2490	1947	78.2
X02	2410	2007	83.3

4.6. **Practical system validation.** Based on the designed original tobacco grading method incorporating multi-attention fine-grained and color texture, an intelligent tobacco sorting and grading system was developed, as shown in Figure 6. An image acquisition device was added to the 6-2 part of the tobacco sorting system in Qinhuangdao, and then the following data were transferred to the tobacco classification system deploying the algorithm designed in this paper to verify the performance of the system. The same grade of tobacco used in Table 5 is used, and the system classification results are shown in Table 7. It can be seen that the tobacco grade classification system based on the method proposed in this paper can accurately classify the tobacco, with high sorting accuracy and stability effectively overcomes the grading errors in the traditional method due to factors such as light, and tobacco morphology, and is able to deal with different states of tobacco samples, which further enhances the intelligent and automated level of tobacco sorting.

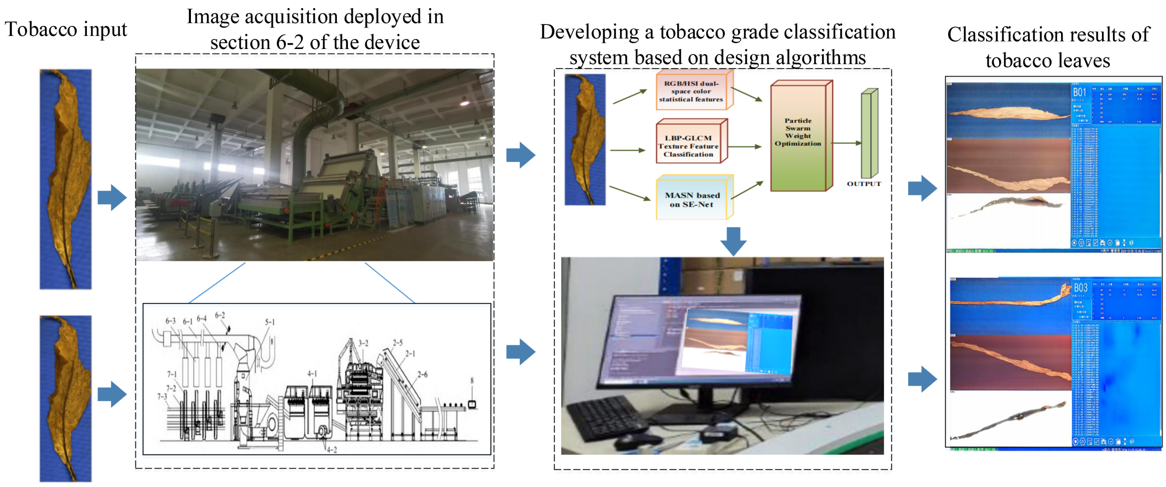


FIGURE 6. Attention sampling moudle

5. **Conclusions.** To improve the accuracy and automation level of tobacco leaf grading, this article proposes an intelligent grading method that integrates multi-attention fine-grained features, color, and texture. This method extracts the color features of tobacco leaves through the RGB/HSI dual-space color statistics strategy, and combines the LBP and GLCM to extract texture features, thereby realizing comprehensive capture of the key visual information of tobacco leaves. Meanwhile, the designed fine-grained classification

Table 7. Recognition rate results of different tobacco ranks

Sample	Sample quantity	Correctly recognized quantity	Recognition rate %
B01	3400	2601	76.5
C01	2800	2112	75.4
C03	3125	2289	73.2
F03	2490	2016	80.9
X02	2410	2100	87.1

network model based on multi-attention sampling effectively makes up for the deficiency of traditional convolutional networks in distinguishing subtle inter-class differences. In the feature fusion stage, the PSO algorithm is introduced to optimize the feature weight allocation, which not only effectively improves the grading accuracy but also solves the classification deviation problem caused by unbalanced weights in traditional fusion methods. Experimental results fully verify the effectiveness of the grading method designed in this paper. The intelligent tobacco leaf classification and grading system designed based on this method realizes the full-process automation of tobacco leaves from image acquisition and feature extraction to grading determination, providing support for the realization of fully automated tobacco leaf classification and grading.

Future research can focus on further optimizing the performance of the algorithm, expanding the size of the dataset, and improving the real-time processing capability of the system to further enhance the sorting efficiency and accuracy.

Acknowledgment. This work was supported by Special Project of the Science and Technology Plan of China National Tobacco Corporation Yunnan Company under Grant No. 2024530000241026.

REFERENCES

[1] B. Huang, Z. Fan, F. Wang, Y. Jiang, G. Xiao and Y. Wen, Intelligent grading of flue-cured tobacco based on VGG16-DenseNet integrated model, *Chinese Tobacco Science*, vol.45, no.3, pp.102-112, 2024.

[2] Y. Tanaka, Near-infrared spectroscopy assisted tobacco grading system, *Journal of Near Infrared Spectroscopy*, vol.18, no.3, pp.211-219, 2010.

[3] F. Muller, Modular design of tobacco grading equipment for multi-variety adaptation, *Biosystems Engineering*, vol.172, pp.89-98, 2018.

[4] K. Liu, Z. Xiong, K. Zhu et al., Historical review and discussion of flue-cured tobacco production in China, *Journal of Kunming University*, vol.33, no.3, pp.48-52, 2011.

[5] Standardization Administration of China, **GB/T 18771.1-2002* Tobacco Vocabulary – Part 1: Tobacco Cultivation, Curing and Grading*, 2002.

[6] Z. Zhuang, *Research on Automatic Tobacco Leaf Grading Method Based on Machine Vision*, Master Thesis, Southwest University, Chongqing, China, 2016.

[7] L. Ma, S. Jiang, B. Wang et al., Design and experiment of intelligent tobacco grading system, *Modern Agricultural Science and Technology*, pp.140-145, 2023.

[8] Y. Liu, J. Long, H. Sun et al., Preliminary study on application mode of intelligent tobacco grading equipment, *Agricultural Development Equipments*, vol.6, 2021.

[9] H. Li, *Tobacco Quality Analysis and Grade Prediction Model Based on Linear Regression and SVM*, Master Thesis, Kunming University of Science and Technology, Kunming, China, 2013.

[10] H. Li, *Research on Tobacco Grading Based on Clustering and Weighted K-Nearest Neighbors*, Master Thesis, Zhengzhou University, Zhengzhou, China, 2017.

[11] C. Xue, X. Cai, J. Song et al., Regional classification of flue-cured tobacco leaves based on principal component analysis and cluster analysis, *Tobacco Science Technology*, vol.51, no.6, pp.34-41, 2018.

[12] X. Yao, F. He, A. Ping, H. Luo and Q. Guan, Leaf tobacco grading method based on PCA-GA-SVM, *Tobacco Science and Technology*, vol.51, no.12, pp.98-105, 2018.

[13] E.-H. Kim, Z. Wang, H. Zong et al., Design of tobacco leaves classifier through fuzzy clustering-based neural networks with multiple histogram analyses of images, *IEEE Transactions on Industrial Informatics*, vol.20, no.3, pp.4698-4709, 2024.

[14] J. Bin, J. Zhou, W. Fan et al., Automatic grading of flue-cured tobacco leaves based on NIR technology and ELM algorithm, *Acta Tabacaria Sinica*, vol.23, no.2, pp.60-68, 2017.

[15] Q. He, X. Zhang, J. Hu, Z. Sheng, Q. Li, S. Y. Cao and H. L. Shen, CMENet: A cross-modal enhancement network for tobacco leaf grading, *IEEE Access*, vol.11, pp.109201-109212, 2023.

[16] S. Li, X. Pan, X. Chen, J. Zhu, B. Wu and Y. Wen, Comparison of tobacco grading methods based on hyperspectral information, *Tobacco Science and Technology*, vol.54, no.10, pp.82-91, 2021.

[17] H. Luo, *Research on Flue-Cured Tobacco Sorting Algorithm Based on Deep Learning*, Master Thesis, Beijing Jiaotong University, Beijing, China, 2023.

[18] J. Zhang, M. Lu, W. Huang, X. Shi and Y. Wang, DSU-Net: A dynamic stage unfolding network for high-noise image compressive sensing denoising, *Neurocomputing*, vol.618, 129071, 2025.

[19] S. Wang, *Research on Quality Grading of Flue-Cured Tobacco Leaves Based on Convolutional Neural Network (CNN)*, Master Thesis, Yunnan Normal University, Kunming, China, 2020.

[20] M. Lu, Q. Zhou, S. Jiang et al., Flue cured tobacco leaf grading method based on deep learning and multi scale feature fusion, *Journal of Chinese Agricultural Mechanization*, vol.43, no.1, pp.158-166, 2022.

[21] J. Zhang, X. Wang, Y. Li et al., An adaptive dual-weighted feature network for insulator detection in transmission lines, *Neural Computing and Applications*, vol.37, no.10, pp.7067-7087, 2025.

[22] R. Wang, X. Liao and X. Chen, Blood glucose prediction based on ensemble learning fusion model, *J. of Medical Informatics*, 2019.

[23] D. Liu, L. Xu, G. Shao et al., Application of particle swarm optimization algorithm in stable structure optimization of Pt-Pd alloy nanoparticles, *Journal of Xiamen University (Natural Science Edition)*, vol.54, no.1, pp.87-92, 2015.

[24] H. Ye, J. Guo and X. Li, Task migration for Cloudlet federation based on improved particle swarm optimization algorithm, *International Journal of Innovative Computing, Information and Control*, vol.20, no.3, pp.693-707, 2024.

[25] J. Zhang, X. Wang, L. Li et al., Parameter recovery of neural network-based Hammerstein system via immersion and invariance adaptive optimization scheme, *Expert Systems with Applications*, vol.274, 127069, 2025.

[26] D. P. Kingma and J. Ba, Adam: A method for stochastic optimization, *arXiv Preprint*, arXiv: 1412. 6980, 2014.

[27] A. A. Ilham, E. Warni and Pahrul, Soft voting classifier with optimized weight using particle swarm optimization on sentiment analysis for online credit and loan application reviews, *International Journal of Innovative Computing, Information and Control*, vol.20, no.2, pp.359-372, 2024.

[28] D. Sun, K. Zhang, H. Zhong et al., Efficient tobacco pest detection in complex environments using an enhanced YOLOv8 model, *Agriculture*, vol.14, no.3, 353, 2024.

[29] J. Lin, D. Yu, R. Pan et al., Improved YOLOX-Tiny network for detection of tobacco brown spot disease, *Frontiers in Plant Science*, vol.14, 1135105, 2023.

[30] F. Li and T. Li, Economic modeling and simulation in the industrial supply chain model using particle swarm optimization algorithm, *International Journal of Innovative Computing, Information and Control*, vol.20, no.1, pp.163-179, 2024.

Author Biography

Mingdeng Fan received bachelor's degree in 2008 from the School of Mechanical and Electrical Engineering at Wuhan University of Technology, China, majoring in Process Equipment and Control Engineering. He is currently working at the Fujian Longyan Golden Leaf Double Roasting Co., Ltd., China. His main research focus is on the technology of leaf processing and re-drying equipment.

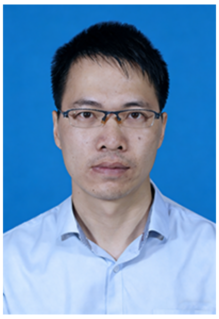
Dayong Xu received master's degree in 2008 from Southeast University, China. He is currently working at the Key Laboratory of Tobacco Processing, Zhengzhou Tobacco Research Institute of CNTC, China. His main research focuses are tobacco processing technology and hyperspectral image processing.

Huajie Li received bachelor's degree in 2004 from the School of Food Science and Engineering at Zhengzhou University of Light Industry, China. He is currently working at the Technical Center of Fujian China Tobacco Industry Corporation, China, primarily engaged in research on cigarette processing technology and equipment.

Xinfeng Li received bachelor's degree in 2011 from the China Central Radio and Television University, China, with a major in Public Administration. He is currently working at the Fujian Longyan Golden Leaf Double Roasting Co., Ltd., China. His main research focus is on process management.

Guojin Xu received bachelor's degree in 2010 from the School of Food and Bioengineering at Zhengzhou University of Light Industry, China, majoring in Tobacco Engineering. He is currently working at the Fujian Longyan Golden Leaf Double Roasting Co., Ltd., China. His main research focus is on innovation in leaf processing and re-drying technology.

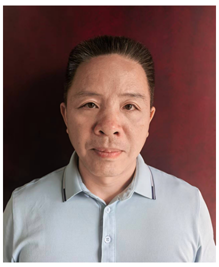
Dengyan Luo reveived bachelor's degree in 2007 from the School of Food Science and Bioengineering at Zhengzhou University of Light Industry, China. He is currently working at the Technical Center of Fujian China Tobacco Industry Corporation, China, specializing in the field of cigarette processing technology.



Danqiang Fu received bachelor's degree in 2019 from Chengyi College of Jimei University, China, majoring in Automation. He is currently working at Fujian Longyan Golden Leaf Double Roasting Co., Ltd., China, primarily responsible for the management of electrical equipment and the construction of automated logistics systems.

Chaochao Wang received bachelor's degree in 2007 from the Tobacco College of Yunnan Agricultural University, China. He is currently working at the Raw Material Research Department of the Technical Center of Fujian China Tobacco Industry Corporation, China. His main research focus is on the production, cleaning, and homogenization processing of cigarette raw materials.

Chen Chen received bachelor's degree in 2022 from Zhengzhou University of Aeronautics, China. He is currently pursuing graduate studies at Zhengzhou University of Light Industry, China, with a primary research focus on deep learning and computer vision. His current work includes the development and optimization of convolutional neural networks for image classification and object detection tasks.



Honghai Wu received bachelor's degree in 2009 from Fuzhou University, China, majoring in Electrical Engineering and Automation. He is currently working at Fujian Longyan Golden Leaf Double Roasting Co., Ltd., China, primarily responsible for production, equipment, and safety management.