Journal of Combinatorial Mathematics
and Combinatorial Computing
www.combinatorialpress.com/jcmcc

Combinatorial Press

Design of a real-time monitoring system for blending ratio and purity of tobacco for hyperspectral imaging

Hongliang Zhang[1,✉], Mei Chen[1], Xiangrong Chen[1], Na Ma[1], Xiang Wei[1]

[1] *Baoding Cigarette Factory, Hebei Baisha Tobacco Co., Ltd., Wangdu, Hebei, 123456, China*

ABSTRACT

In recent years, hyperspectral imaging technology has a large application prospect in quality inspection in the tobacco industry. The study is based on near infrared spectroscopy technology and partial least squares regression method to establish mathematical analysis model of tobacco adulteration ratio of four components, such as expanded tobacco, stalked tobacco, large threaded tobacco and small threaded tobacco, and carry out internal and external inspection. At the same time, TLBO algorithm is used in the optimization of ELM tobacco purity grade determination model to realize the design of tobacco purity monitoring method, and then build the real-time monitoring system of tobacco blending ratio and purity. Tobacco with different purity grades were selected for experimental testing and model comparison analysis. The results show that the constructed PLS model can accurately predict the adulteration content of the four components in tobacco, and the correlation coefficients between the predicted and actual values are above 0.95 ($p < 0.01$), and the relative deviation of the prediction is below 3%.The accuracy of the TLBO-ELM model for identifying tobacco with different grades of purity is 88%, and the classification accuracy in the validation set is improved by 9.32% compared with the ELM model, which is within the acceptable range. It shows a better classification effect than PLS-DA in an acceptable range, which proves that the proposed method can be used for discriminating and monitoring the purity of tobacco. The monitoring system in this paper can be used in the analysis of tobacco blending ratio and purity detection.

Keywords: hyperspectral imaging, partial least squares regression, TLBO algorithm, ELM model, tobacco blending ratio

1. Introduction

During the processing of tobacco raw materials, leaf vein tissue accounts for about 1/4 of the total weight of the tobacco leaf, which is processed by leaf beating and de-stemming to form the raw material of tobacco stalks, and then made into stalks by the filament making process, which

✉ Corresponding author.
E-mail address: baoding789@163.com (H. Zhang).
Received 05 March 2024; Revised 10 May 2024; Accepted 05 December 2024; Published Online 15 April 2025.
DOI: 10.61091/jcmcc127a-495

ultimately becomes an indispensable and important functional module in the tobacco formulation [1-2]. Different components of the tobacco formula are blended at the feeding and blending stage according to the set overall blending ratio, however, the differences in particle size and transport characteristics between different components can lead to component segregation, which makes the blending ratio of each component fluctuate continuously at different spatial locations in the processing time sequence, thus affecting the consistency of product quality, so it is crucial to realize the real-time blending ratio detection for the formulated filaments in the batches[3-6].

In the silk production line, four-filament blending is an important part of the silk production process, and the reasonableness of its process setting and advanced program control directly affect the accuracy of process control [7-8]. According to the cigarette process specification, the four filament blending proportioning accuracy <1.0%, the smaller the value, the higher the accuracy, the more can reflect the advancement of the proportioning system [9-10]. As the most important equipment of the proportioning system, the precision and stability of the belt scale operation, determines the uniformity of the material transportation process, and ultimately affect the sensory quality of the finished product and the intrinsic quality of tobacco [11-12].

However, the current industrial production process still relies more on offline manual picking and weighing to detect the proportion of tobacco adulteration, and is unable to accurately measure the proportion of stems in the finished tobacco in real time [13-14]. In recent years, some researchers have carried out research on the identification of the proportion of tobacco adulteration in tobacco based on image analysis method and spectral analysis method, which mainly involves machine vision technology, deep learning field, near-infrared spectral analysis and hyperspectral analysis, etc., and is characterized by more efficient, more accurate and more rapid [15-17].

In the article, a certain amount of expanded tobacco, stalked tobacco, large threaded tobacco, and small threaded tobacco were selected to be blended according to different ratios, and the tobacco powder was scanned by near-infrared spectroscopy, and a near-infrared spectroscopic identification model was constructed by using partial least-squares regression to determine the blending ratio of the tobacco. After screening the characteristic wavelengths of the near-infrared spectra of the tobacco samples, internal cross-checking and external independent validation of the model were carried out to analyze the correlation between the predicted values and the actual values of different constituents of the tobacco and the prediction errors, and to investigate the prediction effect of the constructed method on the adulteration ratio of the tobacco. Subsequently, to address the problems of poor characterization ability of Extreme Learning Machine and partial failure of hidden nodes, the teaching-learning algorithm is used for optimization, and the TLBO-ELM-based method is proposed to monitor the purity of tobacco filaments. On the basis of the design of the method of tobacco blending ratio and the determination method of purity grade, the real-time monitoring system of tobacco blending ratio and purity is constructed, and the structure and function model of the system are elucidated. Finally, different purity grades of upper, middle and lower tobacco were collected for experimental testing and validation, and the effectiveness of the TLBO-ELM model in monitoring the purity of tobacco was investigated through the comparative analysis between the ELM model and the traditional PLS-DA model, as well as the TLBO-ELM model and the ELM model.

2. Near-infrared spectroscopic identification of tobacco blending ratios

2.1. Materials and Methods

2.1.1. Sample collection. Selected a company silk production line site to take the cigarette big line leaf silk, small line leaf silk, expansion of tobacco and terrier silk 15kg each, which will be placed in the oven, 40 ℃ exhaust baking 2h, and then ground through a 40-mesh sieve, to control the moisture content of 6% to 10% between.

According to the mixing ratio in the actual production process, a total of 250 powders of large

thread leaf silk, small thread leaf silk, expanding tobacco silk and terrier silk were mixed manually and mixed evenly, of which the proportion of large thread leaf silk was 40-50%, the proportion of small thread leaf silk was 15-35%, the proportion of expanding tobacco silk was 3-10%, and the proportion of terrier silk was 4-15%.

2.1.2. Near-infrared spectral scanning. The near-infrared spectra of the samples were collected using the diffuse reflectance module of a Nicolet Antaris Ⅱ Fourier transform near-infrared spectrometer. The parameters were set as follows: spectral scanning range of 10000-4000 cm^{-1}, resolution of 8 cm^{-1}, and scanning times of 60. Before sample collection, the instrument was turned on to warm up for 2 h. Background spectra were collected automatically before each sample collection.

2.1.3. Spectral pre-processing. In order to eliminate the NIR scattering from the inhomogeneous granularity size of the tobacco powder samples, the spectra were preprocessed using multiple scattering correction (MSC). Noise and baseline drift in the NIR spectral information were filtered using Karl Norris filtering and second-order derivative processing. The above spectral preprocessing process was done using TQ analyst software.

2.1.4. PLS model. Partial least squares regression (PLS) was used to develop a mathematical model for the determination of the proportion of tobacco adulteration on a near-infrared analyzer.

Partial least squares regression is a new type of multivariate statistical data analysis method, which is often used for the estimation of model parameters. The partial least squares regression method better solves many problems that are difficult to be solved by ordinary multiple linear regression methods, and it has unique advantages in dealing with the problems of small sample capacity and serious multicollinearity among variables.

There are q dependent variable $\{y_1, y_2, \cdots, y_q\}$ and p independent variables $\{x_1, x_2, \cdots, x_p\}$, and n sample points are observed, thus constituting data tables $X = [x_1, x_2, \cdots x_p]_{n\times p}$ and $Y = [y_1, y_2, \cdots, y_q]_{n\times q}$ for the system of independent variables and the system of dependent variables. Without loss of generality, both sets X and Y are standardized. The partial least squares regression method starts by extracting a principal component t_1 in the independent variable system X and a principal component u_1 in the dependent variable system Y, i.e., t_1 is a linear combination of $x_1, x_2, \cdots x_p$ and u_1 is a linear combination of $y_1, y_2, \cdots, y_q$. These two principal components are extracted with the following two conditions: (1) t_1 and u_1 should carry as much information as possible about the variances in their respective data tables. (2) t_1 and u_1 should be maximally correlated. This indicates that t_1 and u_1 represent the data tables X and Y as well as possible while the component t_1 of the independent variable has the strongest interpretability for the component u_1 of the dependent variable. After the first components t_1 and u_1 have been extracted, partial least squares regression is implemented for X vs. t_1 and Y vs. t_1, respectively. If the regression model has achieved satisfactory accuracy, the algorithm is terminated. Otherwise, a second round of component extraction will be performed using the residual information after X has been interpreted by t_1 and after Y has been interpreted by t_1 to obtain the second principal components t_2 and u_2, and so on until a satisfactory accuracy is achieved. If m components $t_1, t_2, \cdots, t_m$ are finally extracted for X, the partial least squares regression method will implement a regression of $y_k (k = 1, 2, \cdots, q)$ on $t_1, t_2, \cdots, t_m$, which is then reduced to y_k a regression equation on the original independent variable $x_1, x_2, \cdots x_p$

The purpose of partial least squares regression is to seek a combination of components and use it to build a valid linear model for all dependent variables Y. The form of the model is as follows:

$$\hat{y}_k = \beta_{k0} + \beta_{k1}t_1 + \beta_{k2}t_2 + \cdots + \beta_{km}t_m \quad (k = 1,2,\cdots,q) \tag{1}$$

where each component $t_1, t_2, \ldots, t_m$ is a linear combination of independent variables X, and each model has the same component $t_1, t_2, \cdots, t_m$, only their regression coefficients are different.

The specific steps of the algorithm are as follows:

The independent variable $X = \left[x_1, x_2, \cdots x_p\right]_{n\times p}$ and dependent variable $Y = \left[y_1, y_2, \cdots, y_q\right]_{n\times q}$ were standardized separately to obtain the standardized independent variable matrix $E_0 = (E_{01}, E_{02}, \cdots E_{0p})_{n\times p}$ and dependent variable matrix $F_0 = (F_{01}, F_{02}, \cdots, F_{0q})_{n\times q}$.

Denote the first component of E_0 as t_1, $t_1 = E_0 w_1$, w_1 as the first axis of E_0 and $\| w_i \| = 1$. Denote the first component of F_0 as u_1, $u_1 = F_0 c_1$, c_1 as the first axis of F_0 and $\| c_1 \| = 1$.

For t_1, u_1 to represent the maximum variation information of the data in X and Y, respectively, we have:

$$Var(t_1) \to \max \tag{2}$$

$$Var(u_1) \to \max \tag{3}$$

It is also required that t_1 has the maximum explanatory power for u_1, i.e., t_1 and u_1 have to be maximally correlated:

$$r(t_1, u_1) \to \max \tag{4}$$

Combining the above conditions, then the covariance between t_1 and u_1 is required to be maximized, i.e:

$$Cov(t_1, u_1) = \sqrt{Var(t_1)Var(u_1)}\, r(t_1, u_1) \to \max \tag{5}$$

Thus, it can be transformed into the following optimization problem to be solved:

$$\begin{aligned} &\max_{w_1, c_1} \left\langle E_0 w_1, F_0 c_1 \right\rangle \\ &s.t \begin{cases} w_1^{'} w_1 = 1 \\ c_1^{'} c_1 = 1 \end{cases} \end{aligned} \tag{6}$$

Therefore, the maximum value of $w_1^{'} E_0^{'} F_0 c_1$ will be found under the constraints of $\left\|w_1\right\|^2 = 1$ and $\left\|c_1\right\|^2 = 1$. Using the Lagrangian algorithm, this can be obtained:

$$\begin{aligned} E_0^{'} F_0 F_0^{'} E_0 w_1 &= \theta_1^2 w_1 \\ F_0^{'} E_0 E_0^{'} F_0 c_1 &= \theta_1^2 c_1 \end{aligned} \tag{7}$$

$\theta_1 = w_1^{'} E_0^{'} F_0 c_1$, is the value of the objective function of the optimization problem, so w_1 and c_1 are unit eigenvectors corresponding to the largest eigenvalue θ_1^2 of matrices $E_0^{'} F_0 F_0^{'} E_0$ and $F_0^{'} E_0 E_0^{'} F_0$, respectively.

This leads to the component: $t_1 = E_0 w_1, u_1 = F_0 c_1$.

The regressions of E_0 and F_0 on t_1 are then implemented separately: $E_0 = t_1 p_1^{'} + E_1, F_0 = t_1 r_1^{'} + F_1$.

where p_1, r_1 is the regression coefficient and $p_1 = \dfrac{E_0^{'} t_1}{\left\|t_1\right\|^2}, r_1 = \dfrac{F_0^{'} t_1}{\left\|t_1\right\|^2}$.

Denote the residual matrices $E_1 = E_0 - t_1 p_1^{'}$, $F_1 = F_0 - t_1 r_1^{'}$. Replace E_0 and F_0 with the residual matrices E_1 and F_1, and solve for the second axes w_2 and c_2 and the second components t_2 and u_2: $t_1 = E_0 w_1, u_1 = F_0 c_1, \theta_2 = w_2^{'} E_1^{'} F_1 c_2$.

Similarly, w_2 and c_2 are the unit eigenvectors corresponding to the largest eigenvalue θ_2^2 of matrices $E_1^{'} F_1 F_1^{'} E_1$ and $F_1 E_1 E_1 F_1$, respectively.

Calculate the regression coefficient: $p_2 = \frac{E_1^{'} t_2}{\|t_2\|^2}, r_2 = \frac{F_1^{'} t_2}{\|t_2\|^2}$, and the regression equation is obtained:

$$\begin{aligned} E_1 &= t_2 p_{2'} + E_2 \\ F_1 &= t_2 r_{2'} + F_2 \end{aligned} \tag{8}$$

As this calculation continues, if the rank of X is A, then we have:

$$E_0 = t_1 p_1^{'} + t_2 p_2^{'} + \ldots + t_A p_A^{'} \tag{9}$$

$$F_0 = t_1 r_1^{'} + t_2 r_2^{'} + \ldots + t_A r_A^{'} + F_A \tag{10}$$

Since $t_1, t_2, \cdots, t_A$ can all be expressed as a linear combination of $E_{01}, E_{02}, \cdots E_{0p}$, the above equation reduces to a regression equation of $y_k^* = F_{0k}$ on $x_j^* = E_{0j}$, i.e., $y_k^* = \alpha_{k1} x_1^* + \alpha_{k2} x_2^* + \ldots + \alpha_{kp} x_p^* + F_{Ak}$, $k = 1, 2, \cdots, q$.

2.2. Results and Discussion

2.2.1. Raw Spectral Spectra. The NIR spectra of the tobacco blending samples are shown in Figure 1. NIR spectra of weak signals, spectral bandwidth, overlap serious, absorbance difference is small, with complexity and variability of the characteristics of each absorption peak may be a number of different fundamental frequency of the combination of the octave and the combined frequency, there is no sharp peaks and baseline separation of the spectral peaks, a large number of overlapping peaks, the spectral bands belonging to the difficult to determine, and can not be directly used to distinguish between the characteristic absorption peaks, 250 samples of the NIR spectra of samples of the near infrared spectra The NIR spectra of 250 samples were basically consistent with the absorption characteristics of the samples. Therefore, it is indispensable to utilize chemometrics to pre-process the NIR spectra to extract the information characteristics of NIR.

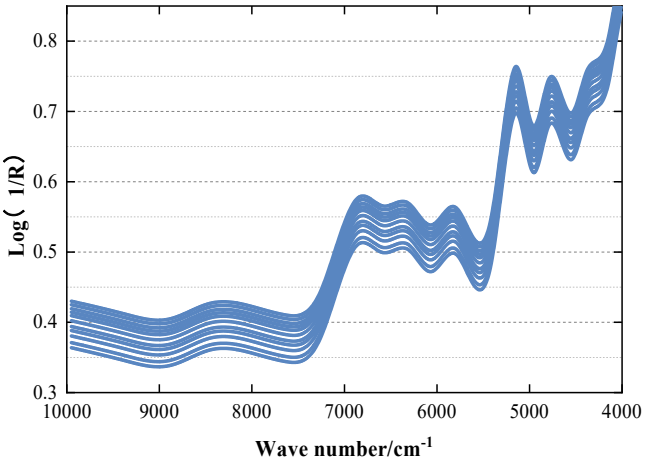


Fig. 1. The near-infrared spectrum of the tobacco wire mixed with the sample

2.2.2. Screening of characteristic wavelengths. The VIP value in the PLS model reflects the contribution of the X variable (spectral absorbance) to the explanatory variable Y (categorical variable), and the larger the value, the greater the contribution to the categorization. The VIP values at different wavelength points in the established PLS model are shown in Fig. 2. In order to effectively exclude the interference of information variables unrelated to classification in the wavelength variables of NIR spectra on the stability and accuracy of the established partial least squares regression model, the wavelength variables with VIP values greater than 1 were screened for regression analysis. The spectral wavelengths with VIP values greater than 1 were screened in the range of 4000-4178, 4725-4834, and 4956-5384 cm^{-1}.

The variance spectrum can reflect the variability of the spectra to some extent. The average spectrum of the tobacco blending samples was calculated, and then the variance spectrum of the average spectrum of the samples was calculated, and the resulting variance spectrum is shown in Figure 3. The wavelengths with larger variance values in the variance spectra were 4000-6200 and 6800-7600 cm^{-1}.

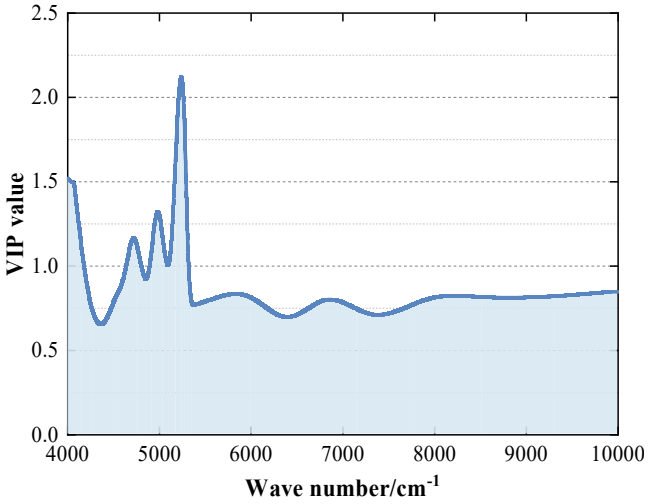


Fig. 2. The VIP values numbered by different waves in the PLS model

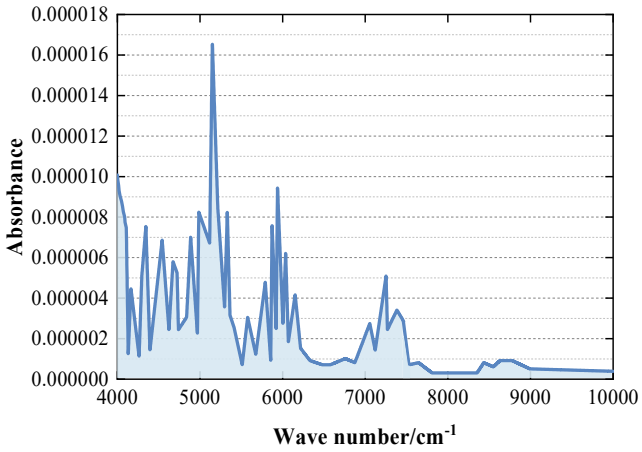


Fig. 3. The spectral variance of the sample

2.2.3. Quantitative analysis modeling. The PLS method is applied to combine the spectral data of the modeled specimen with the corresponding chemical composition content data to optimize the model step by step to reach the best state. The most commonly used methods to test the model are internal cross-checking and external independent validation. The partial least squares method was

used to establish the relationship model between the NIR spectra and the large thread leaf filaments, small thread leaf filaments, stalk filaments and expanded tobacco filaments, and the technical indicators related to the calibration models obtained are shown in Table 1. The correlations of the models were all greater than 0.95, and the cross-validation standard deviation (RMSECV) and prediction standard deviation (RMSEP) were less than 0.2, and the predicted mean relative deviation (RADP) was below 3.4%.

Table 1. The relevant technical indicators of the model

Models	Big line tobacco	Small line tobacco	Swell tobacco	Stalk filament
Principal number	10	7	9	5
R	0.958	0.987	0.988	0.963
RMSECV	0.011	0.015	0.014	0.009
RMSEP	0.016	0.011	0.009	0.007
RADP(%)	2.961	3.365	2.597	3.154
Prediction range	40~50%	15~35%	3~10%	4~15%

In this experiment, 10 samples were produced separately as an independent calibration set, and the constructed model was used to predict the content of the samples' small thread leaf filaments, large thread leaf filaments, stalk filaments, and expanded tobacco filaments. The results of the external validation experiments of the model are shown in Table 2, ** indicates that the correlation coefficients are significant at the 0.01 level, and MRE indicates the mean relative deviation. The correlation coefficients R of the predicted and actual values were 0.9725, 0.9858, 0.9614, and 0.9573, respectively, and the correlations all reached the highly significant level ($p<0.01$). The predicted and actual values were shown by paired t-test ($a=5\%$) that the results were not significantly different ($p>0.05$). Therefore, the prediction model established for the content of small thread leaf filament, large thread leaf filament, stalk filament and expanded tobacco filament can be fully applied to the prediction of the content of tobacco filament blending ratio, so that the uniformity of tobacco filament blending can be grasped.

Table 2. Experimental results of external validation of the model

Sample	Big line tobacco		Small line tobacco		Swell tobacco		Stalk filament	
	Actual /%	Predict /%	Actual /%	Predict /%	Actual /%	Predict /%	Actual /%	Predict /%
1	30.78	25.56	33.28	28.62	34.09	33.24	26.99	34.12
2	20.45	34.04	22.86	30.42	31.48	20.45	22.77	33.81
3	28.50	34.36	27.82	22.72	22.84	33.49	29.60	27.68
4	28.87	24.40	28.99	21.77	29.70	32.79	20.34	33.91
5	31.50	24.13	34.78	32.57	26.56	31.40	20.44	24.43
6	20.25	23.67	24.27	27.06	23.49	27.14	34.38	23.11
7	21.58	23.80	33.23	23.33	30.54	24.87	24.54	26.44
8	26.33	26.09	23.52	22.92	23.24	21.74	27.15	33.93
9	23.63	29.87	21.78	33.09	25.64	33.42	27.16	26.00
10	23.12	29.11	32.69	23.59	22.07	33.06	32.10	30.65
MRE/%	2.00		1.71		2.19		2.86	
R (Double tail)	P<0.01, 0.9725**		P<0.01, 0.9858**		P<0.01, 0.9614**		P<0.01, 0.9573**	
P(T <=t)(Double tail)	0.5324		0.2471		0.6358		0.4686	

3. Design of real-time monitoring system for blending ratio and purity of cigarettes

The purity of tobacco is an important part of its quality evaluation. After recognizing the doping ratio of tobacco using near-infrared spectroscopy, this chapter establishes a monitoring model of tobacco purity based on hyperspectral imaging technology using an extreme learning machine to determine the grade of the purity of tobacco. On this basis, combining the previous chapter and the identification model of tobacco doping ratio, a real-time monitoring system of tobacco doping ratio and purity is designed.

3.1. Data acquisition

A total of 900 tobacco samples from 10 districts for the years 2020-2023 were collected and categorized according to the purity of tobacco into upper (H), middle (M), and lower (L), 300 samples each, which were divided into calibration, validation, and test sets on a 1:1:1 basis. The tobacco samples were dried, pulverized using a cyclone mill and then sieved through a 60-mesh sieve. Tobacco spectra were collected using a Fourier transform near infrared spectrometer (Antaris Ⅱ) using the integrating sphere diffuse reflectance method, which operates in the wave number range of 10,000 to 4,000 cm^{-1} with a resolution of 8 cm^{-1}. The powdered tobacco was placed in a rotating sample cup, flattened using a sample press, and the sample cup was placed on an anhydrous quartz window with a diameter of 50 mm. Instrument performance was verified by an instrument self-test (ValPro system qualification) before sample spectra were taken. Each sample was scanned 60 times and averaged, with each spectrum containing 1573 wavelength points.

After outlier removal and sample division, in order to reduce the influence of spectral noise as well as baseline drift on the experimental results, this study used the SG smoothing method for spectral preprocessing (first-order derivatives, fifteen-point window widths, and second-order polynomials), and utilized the MC-UVE method for variable screening.

3.2. Research methodology

In this study, the TLBO algorithm is used to optimize the number of nodes in the hidden layer during the ELM operation, so as to improve the correct rate of cigarette classification and achieve the optimization of the TLBO-ELM cigarette purity determination model.

3.2.1. Extreme Learning Machines. The Extreme Learning Machine (ELM) algorithm greatly improves the learning efficiency compared to traditional neural networks.ELM is divided into three layers, including the input layer, the hidden layer and the output layer, because the network has no feedback structure, and the input information is transmitted unidirectionally in one direction, so the learning efficiency is effectively improved.

In practical applications, the number of nodes in the input layer and the number of nodes in the output layer of ELM can be set flexibly according to the number of feature parameters and the number of state categories. Because of its simplicity, convenience, high efficiency, high recognition accuracy, strong generalization and other advantages, ELM has been widely used in problems such as state recognition and classification.

The ELM model is built in two steps, step 1 requires randomly generating the input layer weights w and the implied layer bias b of the implied layer nodes to finally obtain the implied layer output:

$$H(x)=\begin{bmatrix} g(w_1\cdot x_1+b_1) & \cdots & g(w_L\cdot x_1+b_L) \\ \vdots & & \vdots \\ g(w_1\cdot x_N+b_1) & \cdots & g(w_L\cdot x_N+b_L) \end{bmatrix}_{N\times L} \tag{11}$$

Step 2 calculates the output weights of the implicit layer nodes by solving the linear system method by least squares method β. The standard SLFNs are modeled as shown in equation (12):

$$\sum_{i=1}^{N} g(w_i \cdot x_j + b_i)\beta_i = o_j, j = 1,2,\ldots,N \tag{12}$$

where o represents the network output. Since the training objective is to get the minimization error, Eq. (12) can be expressed as:

$$\sum_{j=1}^{N} \| o_j - y_j \| = 0 \tag{13}$$

where y is the actual labeled value. In some cases, the minimization error of the ELM model network is able to approximate any $\varepsilon < 0$, and the network is denoted as:

$$\sum_{i=1}^{N} g(w_i \cdot x_j + b_i)\beta_i = y_j, j = 1,2,\ldots,N \tag{14}$$

When ELM solves for the implicit layer output weight β, the above equation can be simplified as:

$$H\ \beta = Y \tag{15}$$

where H is the output vector of the hidden layer of the neural network. The output weights β can be found by the following equation:

$$\beta = H^{+}Y \tag{16}$$

where H^{+} is the generalized inverse matrix of H.

From the above two steps of the ELM model building process, it can be concluded that ELM is very different from other gradient learning algorithms, the initial weights and bias do not need to be repeatedly adjusted, and the solution is very direct. Comparing ELM with other traditional neural networks, the learning time of the ELM model is fast, based on the principle of least squares, which eliminates the problem of falling into the local optimum.ELM has certain advantages over back propagation neural networks, but the characterization ability of the single-layer network is poor, so it needs an effective indicator when selecting the signal feature parameters in order to make the ELM model play a better classification and recognition effect.ELM's The initial weight matrix parameters are set randomly, and once they are set, they will not change until the end of the program, which may make some parameters zero, resulting in the failure of the hidden part of the node. Now the above problems can be solved by optimizing the parameters, and this chapter is to improve the classification recognition accuracy and efficiency of ELM model by teaching and learning optimization algorithm.

3.2.2. Teaching and learning algorithms. The Teaching and Learning Algorithm (TLBO) is a simulation of a class as a population, in which the overall level of students in the class is improved through the teaching phase of the teacher, and the individual level is further improved through the mutual learning phase of the students, so as to achieve the goal of continuously optimizing the population. Among them, the teacher and the students are the individuals in the algorithm, the teacher is the individual with the best adaptation value, the number of subjects studied by each student is the number of control variables, and the students' performance is the function adaptation value. The specific definition is as follows:

For the maximum optimization problem: $\max\{f(x) \mid x \in S\}$, where $f(x)$ is the fitness function, $S = \left\{x \mid x_j^L \le x_j \le x_j^U, j = 1,2,\cdots d\right\}$ is the search space, $x = (x_1, x_2, \ldots, x_d)$ is the search points in the search space, d is the number of dimensions (i.e., the number of decision variables) of each search

point, x_j^L is the upper bound of each dimension of the search points: x_j^U is the lower bound of each dimension of the search points. The individuals (i.e., search points) in the entire population can be represented as $x^j=(x_1^i,x_2^j,\cdots x_d^i),i=1,2,\cdots NP$, where $x_j^i,j=1,2,\cdots d$ is the decision variable for each search point and NP is the number of search points (i.e., population size).The TLBO algorithm nomenclature is explained as follows:

Class: the set of all search points is called a class (i.e., population).

Student: each individual in the class $x^i=(x_1^i,x_2^i,\cdots,x_d^i),i=1,2,\cdots,NP$.

Teacher: the individual in the class with the best adaptation value, denoted by $x_{teacher}$.

Subjects: each individual for each dimensional variable, i.e. $x=(x_1,x_2,\ldots,x_d)$, d is the number of subjects.

A class can be represented as follows:

$$\begin{bmatrix} x^1 & f(x^1) \\ x^2 & f(x^2) \\ \vdots & \vdots \\ x^{NP} & f(x^{NP}) \end{bmatrix}=\begin{bmatrix} x_1^1 & x_2^1 & \cdots & x_d^1 & f(x^1) \\ x_1^2 & x_2^2 & \cdots & x_d^2 & f(x^2) \\ \vdots & \vdots & & \vdots & \vdots \\ x_1^{NP} & x_2^{NP} & \cdots & x_d^{NP} & f(x^{NP}) \end{bmatrix} \tag{17}$$

where, $x^{'}=(x_1^{'},x_2^{'},\cdots,x_d^{'})$ is the students in the class, NP is the population size, d is the number of subjects studied and $x_{teacher}=\arg\min f(x^{'}),(i=1,2,\ldots,NP)$ is the teacher in the class.

The basic principle of TLBO algorithm is as follows:

1) Population initialization. Let the population size be NP and the number of subjects be d, then the random initial population is:

$$x=\begin{bmatrix} x_1^1 & x_2^1 & \ldots & x_d^1 \\ x_1^2 & x_2^2 & \ldots & x_d^2 \\ \vdots & \vdots & & \vdots \\ x_1^{NP} & x_2^{NP} & \ldots & x_d^{NP} \end{bmatrix} \tag{18}$$

The initial students are generated as shown in equation (19):

$$x_j^i=x_j^L+rand(0,1)\times(x_j^U-x_j^L)i=1,2,\cdots,NP;j=1,2,\cdots d \tag{19}$$

where x^i is the i nd student, x_i^U and x_j^L are the upper and lower bounds of each one-dimensional decision variable for student x^i, and $rand(0,1)$ is a random number between 0 and 1.

2) Teaching stage. The optimal individual in the population is selected as the teacher, and then each student learns according to the difference between the teacher's and the student's mean MEAN in the process shown below:

First, calculate the difference between the teacher and student mean MEAN.

$$difference=r_i^*(x_{teacher}-T_f*mean) \tag{20}$$

where $mean=\frac{1}{NP}*\sum_{NP}^{i=1}x^i$ is the average value of the students. r_i is the learning step, a random number between 0 and 1. $difference$ is the difference between the student mean $mean$ and the teacher $x_{teacher}$. T_f is the teaching factor, $T_f=round(1+rand(0,1))$.

Then, differential learning is carried out by Equation (21):

$$x_{nev}^{i} = x^{i} + difference \tag{21}$$

where x^{i} and x_{new}^{i} are the values of the ird student before and after the instruction, respectively.

The performance of student x^{i} is measured by its adaptation value $f(x^{i})$, where the larger $f(x^{i})$ is the better the performance of x^{i}, and vice versa, the smaller $f(x^{'})$ is the worse the performance of $x^{'}$. Individuals are updated at the end of the instruction and if $f(x_{\text{new}}^{i}) > f(x^{i})$, i.e., x_{new}^{i} performs better, the old member is replaced by the new member x_{new}^{i}, and vice versa, the old member is retained.

3) Learning phase. Select a student x^{i} from the class after teaching and take a different student x^{k} to learn from each other with him and compare the two $f(x^{i})$ using equation (22) for mutual learning:

$$x_{new}^{i} = \begin{cases} x^{i} + r^{*}(x^{i} - x^{k}) & f(x^{i}) > f(x^{k}) \\ x^{i} + r^{*}(x^{k} - x^{i}) & f(x^{i}) < f(x^{k}) \end{cases} \tag{22}$$

where r is generally taken as a random number between 0 and 1 for each student's learning factor.

Then the update of individual students is carried out, if $f(x_{new}^{i}) > f(x^{i})$ i.e. x_{new}^{i} is more excellent in learning, the old member is replaced with the new member x_{new}^{i} and vice versa the old member is retained.

The steps for implementing the teaching and learning algorithm are as follows:

Step 1: Initialize the population parameters, initialize the population according to equation (19), and calculate the adaptation value of individuals of the population.

Step 2: Teaching phase, the teacher's teaching process, find out the optimal adaptation value individual for the teacher, and each student learns from the teacher through formula (21).

Step 3: Learning phase, students' mutual learning process, two different students are randomly selected to communicate and learn through Eq. (22).

Step 4: After the teaching and learning phase is completed, the adaptation value of individuals in the population is calculated, and if the adaptation value of the current individual is better than the adaptation value of the individual in the previous generation, the individual is updated and enters the next iteration.

Step 5: Judge whether the termination condition is satisfied, if so, terminate the iteration, otherwise, loop the execution of steps 2-step 4.

3.3. Real-time monitoring system design

3.3.1. System architecture design. The real-time monitoring system of tobacco blending ratio and purity is based on the main line of data collection in the silk-making workshop and statistical analysis of the production process data. Based on the configuration of the data collection model, the near-infrared spectral recognition model is used to analyze the blending ratio of the tobacco, and then classify the purity grade of the tobacco based on the TLBO-ELM model, and ultimately achieve the smooth flow of the real-time monitoring system of the tobacco blending ratio and purity in this tobacco limited liability company. The system is designed to realize the smooth flow of the tobacco blending ratio and purity real-time monitoring system of the tobacco company and promote the deep integration of the business of various departments.

The structure of the monitoring system is shown in Figure 4, which is mainly divided into four levels:

1) Application platform. Data storage and software platform construction.

2) Data collection model. The data collection model includes the model configuration of the production line of the silk workshop, the model configuration of the production equipment, and the model configuration of the equipment address point.

3) Tobacco blending ratio and purity monitoring. Data collection, blending ratio analysis and purity level monitoring of tobacco.

4) Human-computer interaction. Provide statistical analysis and display interface of business data.

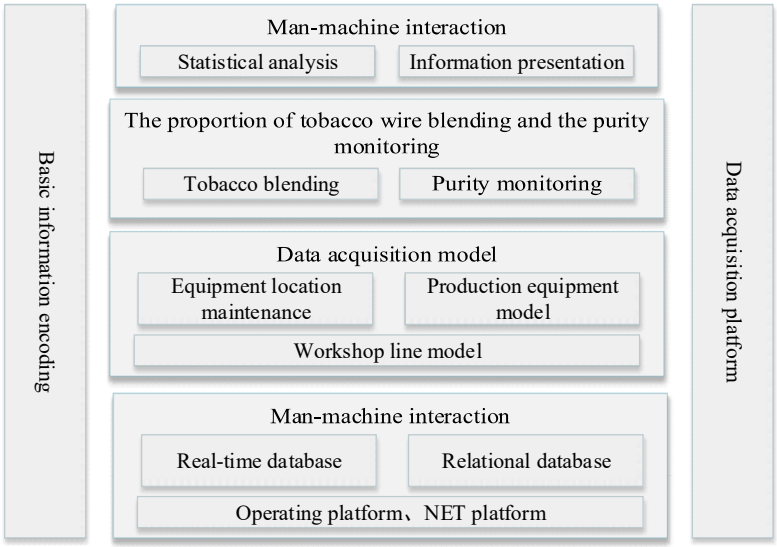


Fig. 4. The structure design of the monitoring system

3.3.2. Functional module design. Tobacco blending ratio and purity real-time monitoring system mainly includes tobacco blending ratio module and tobacco purity real-time monitoring module. Tobacco blending ratio module includes data collection and confirmation of data collection points, data collection model configuration, and analysis of tobacco blending ratio. Tobacco purity real-time monitoring module includes online monitoring and offline monitoring. Thus completing the silk production process management from data collection, production process monitoring to quality control analysis.

1) Tobacco blending ratio module. Tobacco blending ratio module includes data collection and confirmation of data collection points, data collection model configuration, and analysis of tobacco blending ratio. The monitoring system requires each cigarette factory to collect the data collection points of the equipment in each work unit of each production line. The data collection model configuration needs to model each production line, work unit, equipment, collection point and labeling point of the silk workshop through the tree structure, and the model construction of the data collection needs to follow the rules of the software development to depict the structure of the equipment in the silk workshop. Tobacco blending ratio analysis utilizes near-infrared spectroscopy combined with PLS modeling to identify and classify the composition of tobacco samples.

2) Purity real-time monitoring module. The purity real-time monitoring module is divided into online quality module and offline quality module. This part combines the high-high-precision imaging technology and uses the extreme learning machine model optimized by the teaching-learning algorithm to classify the purity of the tobacco samples and realize the real-time monitoring and management of the purity of the tobacco.

4. Results of the monitoring analysis and discussion

Based on the designed real-time monitoring system of tobacco blending ratio and purity, this chapter analyzes the effect of monitoring the purity of tobacco therein through the comparative experiments of the model.

4.1. Comparison of PLS-DA and ELM models

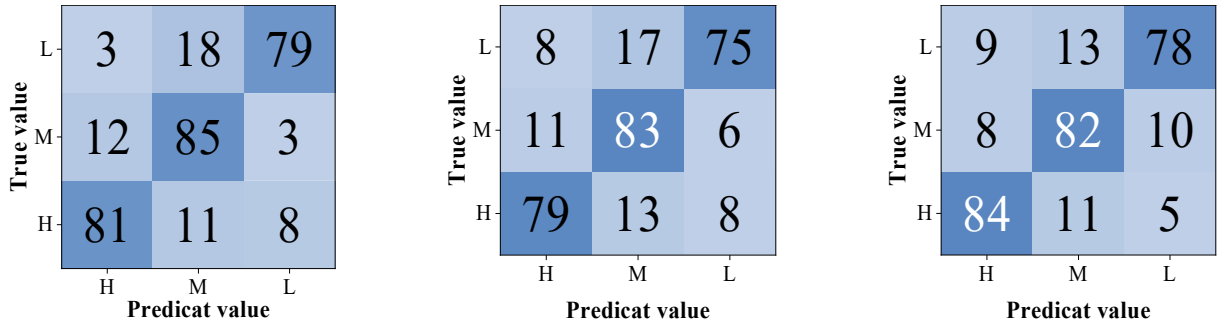
The advantages and application significance of the ELM model were verified by establishing the conventional PLS-DA tobacco purity monitoring model and comparing it with the ELM tobacco purity monitoring model.The optimal number of latent variables for PLS-DA was determined to be 25 based on the ten-fold cross validation (RMSECV=0.3524). The corresponding key variables were screened and calculated using the same spectral preprocessing method and variable screening method as the ELM model.

The results of ELM and PLS-DA tobacco purity monitoring models are compared in Table 3, and the corresponding confusion matrices of the prediction results of the PLS-DA model for the calibration, validation, and test sets are shown in Fig. 5. The correct classification rates of the corrected, validated, and test sets of the ELM tobacco purity monitoring model for the sub-sites were 83.06%, 83.88%, and 86.06%, respectively, which were better than those of the traditional PLS -DA method.

The purity monitoring model established by PLS-DA had a lower correct classification rate for lower purity tobacco, which was mostly misclassified as middle tobacco, reducing the overall correct classification rate. Thus, it is of practical significance to use ELM to establish the tobacco purity monitoring model in this study, based on which the ELM model was optimized by using TLBO, which highlights the advantages of this study.

Table 3. Comparison of the effects of PLS-DA and ELM models

Models	Number of latent variables/hidden layer nodes	Classification accuracy (%)		
		Correcting set	Verification set	Test set
PLS-DA	25	81.67	79.00	81.33
ELM	60	83.06	83.88	86.06



(a)Training set (b)Calibration set (c)Testing set

Fig. 5. Confusion matrix of the prediction results of the PLS-DA model

4.2. TLBO-ELM model analysis

Although the generalization ability of the pre-built ELM model is good, its classification correctness is not good, thus the TLBO algorithm is considered to optimize the number of nodes in its hidden layer. The maximum number of nodes in the hidden layer is set to 200, the ELM activation function

g(x) = 'sig', and the TLBO parameters are set as follows: n Pop=25, MaxIt=55, nVar=2, VarMin=20, VarMax=200. The parameter optimization process is shown in Fig. 6, and the optimal hidden layer is obtained. The number of nodes is 98, the minimum value of adaptation is 1.075, and the classification correctness of the validation set is improved to 93.2%, which is 9.32% compared to the ELM model.

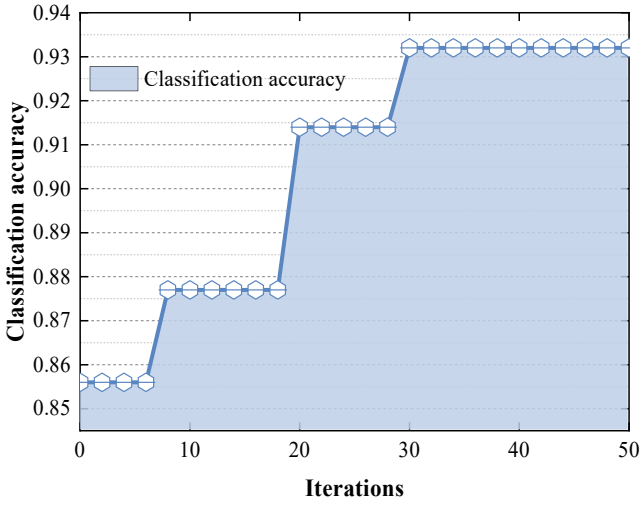


Fig. 6. Parameter optimization process of TLBO-ELM

The model was externally validated using a test set and the classification was correct at 88%, a slight improvement compared to the ELM model, and the confusion matrix results are shown in Fig. 7. The classification accuracies for the upper, middle, and lower purity tobaccos were 91%, 88%, and 85%. Although there were cases of misprediction, they were generally misclassified as neighboring classes and the classification results were within acceptable limits.

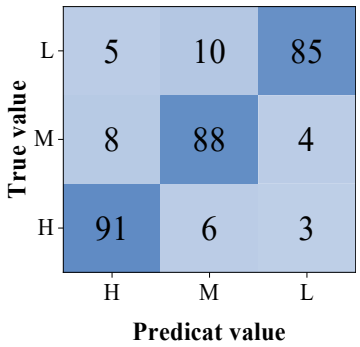


Fig. 7. Confusion matrix of the prediction results about TLBO-ELM model testing set

5. Conclusion

In this paper, the identification model for the analysis of the adulteration ratios of expanded tobacco, stalked tobacco, large threaded tobacco, and small threaded tobacco was firstly established by using near-infrared spectroscopy in combination with partial least squares regression method. The internal cross-test and external validation show that the correlation coefficients of the predicted and actual values of tobacco adulteration are greater than 0.95, and the correlation is presented as significant at the 1% level, and there is no significant difference in the results of the paired t-test, and the identification model established for the content of small thread leaf filament, large thread leaf filament, terrier filament, and swollen tobacco filament can be completely applied to the prediction of the adulteration content of tobacco filaments. Secondly, the teaching-learning algorithm is used to

optimize the extreme learning machine model to construct a real-time monitoring method of tobacco purity, and thus design a monitoring system that contains two functional modules, namely, the analysis of tobacco blending ratio and the monitoring of purity. In the determination of the purity of different grades of tobacco, the classification accuracy of the ELM model is better than that of the traditional PLS-DA method in each dataset, with an accuracy of more than 83%. Compared with the ELM model, the TLBO-ELM model improves the classification correct rate of the validation set from 83.88% to 93.2%, which is an improvement of 9.32%, and the overall accuracy rate reaches 88%, which is able to complete the monitoring and identification of the purity grade of tobacco more accurately. Comprehensive experimental results show that the real-time monitoring model of tobacco blending ratio and purity designed in this paper has good practicability. Hyperspectral imaging technology is applied to the quality and safety detection in the tobacco industry, can be synthesized to get the comprehensive detection information of the internal and external quality of the product, this internal and external quality information both features, so that hyperspectral image technology in the tobacco industry detection has a greater application prospects. At this stage, the use of hyperspectral imaging technology for the detection of the tobacco industry is still in the research and development stage, with the continuous improvement of the spectral resolution, hyperspectral imaging can record more and more rich tobacco quality information.

References

[1] Wang, L., Jia, K., Fu, Y., Xu, X., Fan, L., Wang, Q., ... & Niu, Q. (2023). Overlapped tobacco shred image segmentation and area computation using an improved Mask RCNN network and COT algorithm. Frontiers in Plant Science, 14, 1108560.

[2] Zhu, B., An, H., Li, L., Zhang, H., Lv, J., Hu, W., ... & Li, D. (2024). Characterization of Flavor Profiles of Cigar Tobacco Leaves Grown in China via Headspace–Gas Chromatography–Ion Mobility Spectrometry Coupled with Multivariate Analysis and Sensory Evaluation. ACS omega, 9(14), 15996-16005.

[3] Chen, G., Wei, H., Long, Y., Zhu, B., Zhu, J., Shen, L., ... & Chen, M. (2024, February). Research and Application of Improving the Whole Cut Tobacco Rate in the Mixing Cabinet. In Proceedings of the International Conference on Industrial Design and Environmental Engineering, IDEE 2023, November 24–26, 2023, Zhengzhou, China.

[4] Wu, D., Chen, L., Gu, Q., & Luo, J. (2023, December). Research on Tobacco Foreign Object Detection Based on Deep Learning of Texture Features. In 2023 4th International Conference on Information Science and Education (ICISE-IE) (pp. 1-4). IEEE.

[5] Li, Q., Li, Y. F., Zhang, Y., Chen, Q., Huang, H., Chen, H., ... & Chen, X. D. (2018). Drying kinetics study of irregular fibril materials in a “differential” laboratory rotary dryer: Case study for cut tobacco. Drying Technology, 36(5), 523-536.

[6] Li, Z., Wang, Z., Wang, J., Wu, Z., Zhang, L., & Liu, H. (2023, December). Modeling and simulation analysis of tobacco sorting system and strategy. In 2023 International Conference on Applied Physics and Computing (ICAPC) (pp. 345-348). IEEE.

[7] Sahu, A., & Dante, H. (2018, May). Non-destructive rapid quality control method for tobacco grading using visible near-infrared hyperspectral imaging. In Image Sensing Technologies: Materials, Devices, Systems, and Applications V (Vol. 10656, p. 1065603). SPIE.

[8] Zou, Q., Zhang, T., Zhao, Y., Chen, R., Yang, L., Hu, Q., ... & Feng, H. (2021). Effect of cut tobacco size and distribution on critical cigarette quality characteristics of “slim cigarette” processing technology. In Journal of Physics: Conference Series (Vol. 1748, No. 6, p. 062043). IOP Publishing.

[9] Zhu, W., Liu, H., Zhou, B., Ding, M., Wang, B., & Liu, B. (2022). Nondestructive Detection of Stem Content

in Tobacco Strips Using X-Ray Imaging Analysis. Contributions to Tobacco & Nicotine Research, 31(3), 142-150.

[10] He, Y., Wang, H., Zhu, S., Zeng, T., Zhuang, Z., Zuo, Y., & Zhang, K. (2018). Method for grade identification of tobacco based on machine vision. Transactions of the ASABE, 61(5), 1487-1495.

[11] Niu, Q., Liu, J., Jin, Y., Chen, X., Zhu, W., & Yuan, Q. (2022). Tobacco shred varieties classification using Multi-Scale-X-ResNet network and machine vision. Frontiers in Plant Science, 13, 962664.

[12] Hou, J., Wang, D. D., & Wang, D. (2020, July). Research on adjusting size proportion of cut tobacco. In IOP Conference Series: Materials Science and Engineering (Vol. 892, No. 1, p. 012116). IOP Publishing.

[13] Xin, X., Gong, H., Hu, R., Ding, X., Pang, S., & Che, Y. (2023). Intelligent large-scale flue-cured tobacco grading based on deep densely convolutional network. Scientific Reports, 13(1), 11119.

[14] Feng, X., Pan, A., Ren, Z., Hong, J., Fan, Z., & Tong, Y. (2023). An adaptive dual-population based evolutionary algorithm for industrial cut tobacco drying system. Applied Soft Computing, 144, 110446.

[15] Gendall, P., & Hoek, J. (2024). Regulating flavours and flavour delivery technologies: an analysis of menthol cigarettes and RYO tobacco in Aotearoa New Zealand. Tobacco control, 33(e1), e122-e124.

[16] Chao, M., Kai, C., & Zhiwei, Z. (2020). Research on tobacco foreign body detection device based on machine vision. Transactions of the Institute of Measurement and Control, 42(15), 2857-2871.

[17] Jia, K., Niu, Q., Wang, L., Niu, Y., & Ma, W. (2023). A New Efficient Multi-Object Detection and Size Calculation for Blended Tobacco Shreds Using an Improved YOLOv7 Network and LWC Algorithm. Sensors, 23(20), 8380.