

Monitoring green tea fixation quality by intelligent sensors: comparison of image and spectral information

Yuyu Chen,^{ID} Huiting Wu, Ying Liu, Yujie Wang, Chengye Lu, Tiehan Li,^{ID} Yuming Wei and Jingming Ning^{ID}*



Abstract

BACKGROUND: Intelligent monitoring of fixation quality is a prerequisite for automated green tea processing. To meet the requirements of intelligent monitoring of fixation quality in large-scale production, fast and non-destructive detection means are urgently needed. Here, smartphone-coupled micro near-infrared spectroscopy and a self-built computer vision system were used to perform rapid detection of the fixation quality in green tea processing lines.

RESULTS: Spectral and image information from green tea samples with different fixation degrees were collected at-line by two intelligent monitoring sensors. Competitive adaptive reweighted sampling and correlation analysis were employed to select feature variables from spectral and color information as the target data for modeling, respectively. The developed least squares support vector machine (LS-SVM) model by spectral information and the LS-SVM model by image information achieved the best discriminations of sample fixation degree, with both prediction set accuracies of 100%. Compared to the spectral information, the image information-based support vector regression model performed better in moisture prediction, with a correlation coefficient of prediction of 0.9884 and residual predictive deviation of 6.46.

CONCLUSION: The present study provided a rapid and low-cost means of monitoring fixation quality, and also provided theoretical support and technical guidance for the automation of the green tea fixation process.

© 2022 Society of Chemical Industry.

Supporting information may be found in the online version of this article.

Keywords: green tea; micro-NIR spectroscopy; computer vision system; fixation degree; moisture content; sensory evaluation

INTRODUCTION

As one of the most traditional non-alcoholic beverages worldwide, green tea has a large consumer base.¹ Green tea is a non-fermented tea that undergoes the main processes of fixing, shaping and drying from fresh tea leaves to finished tea products. Fixing is considered to be the most critical process with respect to forming the 'green leaf clear soup' quality appearance of green tea. During the fixation process, fresh tea leaves are treated with high temperatures to terminate oxidase activity, accompanied by a dehydration reaction.² When fixation is insufficient, this leads to high moisture content of fresh leaves, which makes them prone to the appearance of red stems and red leaves during subsequent processing; when fixation is excessive, the surface of fresh leaves loses too much water and swells, causing the leaf surface to blister and break, leaving burst spots and scorched spots. Therefore, the degree of leaf dehydration during the fixing process directly affects the green tea quality.³ Traditionally, tea makers evaluate the fixation degree subjectively according to variations in color, appearance and aroma of a fixed tea sample, which lacks objectivity and accuracy. Moisture content, which is the main standard for judging the degree of fixation, is measured using the traditional drying weighing

method, and is accurate but time-consuming. Traditional evaluation methods fail to fulfill the requirement of real-time monitoring of fixation quality in large-scale production.⁴

Near-infrared spectroscopy (NIRS) and computer vision systems (CVS) have the potential to sense changes in the internal and external properties of tea in-process products in real-time and are important technologies for achieving intelligent tea processing.⁵ NIRS is a mature non-destructive detection technique that has evolved into an indispensable analytical tool for quality detection in the tea processing industry. Compared to the costly and bulky Fourier transform near-infrared spectroscopy (FT-NIRS), the low-cost, flexible micro-NIRS is more suitable for online monitoring.⁶ Shen *et al.*⁷ used micro-NIRS to achieve rapid detection of the moisture of black tea during withering, and Jin *et al.*⁸ determined fermentation quality by analyzing the fermentation degree

* Correspondence to: J Ning, State Key Laboratory of Tea Plant Biology and Utilization, Anhui Agricultural University, Hefei 230036, China. E-mail: ningjm1998009@163.com

State Key Laboratory of Tea Plant Biology and Utilization, Anhui Agricultural University, Hefei, China

of black tea in relation to catechins and theaflavins. However, all of the abovementioned studies focused on the black tea processing process, and few studies have focused on green tea processing. The CVS, comprising an intelligent sensory analysis technology, can visually convey visual attributes and accurately quantify information about the appearance of tea. It assesses the quality of the in-process product by extracting the appearance features of tea. Dong *et al.*⁹ used image color features combined with non-linear methods to quantitatively evaluate the changes in sensory quality and pigment during the fermentation. The same team also established a non-destructive detection method for the moisture of black tea during withering by CVS.¹⁰ Chen *et al.*¹¹ predicted the chemical compositions of green tea before and after fixation based on color information and further improved green tea quality through an adaptive fuzzy logic control system. Regrettably, the fixation process for this study was performed offline.¹¹ In addition, rapid evaluation of green tea fixation quality utilizing CVS has rarely been reported.

Therefore, the present study targeted the green tea automatic production line and analyzed the feasibility of the rapid at-line monitoring of green tea fixation quality using smartphone-coupled micro-NIRS and a self-built CVS. The specific objectives were: (i) collect *in situ* data of fixation stage samples using two sensors; (ii) explore effective data analysis methods to establish classification models for fixation degree and prediction models for moisture content; and (iii) compare the performance of two low-cost sensors suitable for industrial applications in fixation quality monitoring.

MATERIALS AND METHODS

Collection and evaluation of green tea fixation samples

Green tea samples with different fixed temperatures were collected from Lvming Tea Factory (Jinzhai, Anhui Province, China) following the traditional strip green tea process in July 2021. Specifically, fresh leaves [*Camellia sinensis* (L.) var. Baiye 1] with a standard of one bud and two leaves were picked and subsequently transported to the tea factory. The fresh leaves were evenly spread at a temperature of 18.5 °C and a humidity of 80% relative humidity for 4 h and then transferred to a type 6CST-50 roller-hot air fixation machine (Zhejiang Green Peak Machinery Co., Quzhou, Zhejiang Province, China) for fixing. To restore the real trend of blade changes during fixing as much as possible, fixed tea samples were collected at machine cylinder wall temperatures of 210, 240, 270, 300, 330 and 360 °C by controlling leaf throwing and machine speed constant. Specifically, fresh leaf throwing started when the machine barrel wall reached the set temperature, and samples collection started when the leaf state at the leaf outlet was stable. Six fixation temperatures samples were carried out in six batches, with 20 replicate samples taken continuously at each fixation temperature, for a total of 120 samples. All samples were collected *in situ* at the end of conveyor belt. The experimental process is presented in Fig. 1.

For the collected samples, image information and spectral information were obtained at-line using a self-built CVS and a micro-NIRS. Each sample was accurately weighed before being transported back to the laboratory. Accordingly, 100 g of samples were collected from each of six fixation temperatures and dried to approximately 5% moisture for post sensory review. The sensory review is conducted according to the Chinese National Standard GB/T 23776-2018. Details and results of the review are provided in the Supporting information (Table S1).

Image collection and variable extraction

The self-built CVS is affordable; it consists of an industrial camera (MER-500-14U3M/C; Daheng, Beijing, China), a camera lens (M0824-MPW2; Computar, Carlsbad, CA, USA), an arched light source and a carrier platform connected to a computer, with a black dark box made of acrylic panels on the outside. The camera is fixed at the top of the dark box; the camera has the following parameters: 5 MP sensor, focal length of 8 mm, aperture of F2.4, exposure time of 950 µs, externally triggered acquisition, resolution of 2592 × 1944 and vertical distance of 23 cm from the sample to ensure a clear image and a constant distance. Before image collection, the light source in the dark box was switched on 20 min in advance to maintain a stable internal environment. The collected samples were spread evenly on the collection trays, and then placed on the carrier platform for image collection. Spectral collection followed after image collection. The acquired original images of the tea samples were imported into MATLAB2014a (MathWorks Inc., Natick, MA, USA) for processing. Here, we selected the image global as the region of interest to extract the following feature variables. (i) The gray level gradient co-occurrence matrix (GLGCM) was used to extract m (mean), δ (standard deviation), γ (smoothness), μ (third moment), U (uniformity) and e (entropy) as six texture feature variables.¹² (ii) The average pixel information of each color channel of R, G and B is also extracted. Then, the RGB color space is transformed into HSV and CIELAB color space based on a previously reported method.¹³

To eliminate gradient differences between variables of different latitudes, image feature variable data were processed by maximum-minimum normalization. The preprocessed data were in the range of −1 to 1. The maximum-minimum normalization algorithm is based on:

$$x_{\text{norm}} = \frac{x_i - \min(x)}{\max(x) - \min(x)} \quad (1)$$

where x_i is the image pixel point value, and $\min(x)$ and $\max(x)$ are the minimum and maximum values of the image pixels, respectively.

Spectral collection and feature selection

A low-cost micro-NIR spectrometer (NIR-S-R2; InnoSpectra Corporation, Taiwan, China) coupled to a smartphone was used to collect spectral data at wavelengths in the range 900–1700 nm for the samples. The spectrometer offers a resolution of 10–12 nm and sizes 75 × 58 × 26.5 mm³. An Honor 10 GT phone (Huawei Technologies Co., Shenzhen, China) was used to collect and store spectral data. Spectral collection was performed at room temperature (approximately 25 °C, using micro-NIRS in direct contact with the surface of the fixation tea samples). Three scans were taken for each sample to average the spectral data as a representative spectrum of the sample. The obtained spectrum files were stored in .csv format on the smartphone and imported into MATLAB2014a on an external computer via the USB line for data processing. To eliminate the impact of background information and noise from the raw spectra, 228 raw spectral wavelengths were preprocessed with the standard normal variable (SNV), first derivative (1D) and second derivative (2D). SNV preprocessing is used to eliminate errors in the diffuse reflectance spectrum caused by surface scattering and light range variations of the sample being measured. The derivative correction can effectively

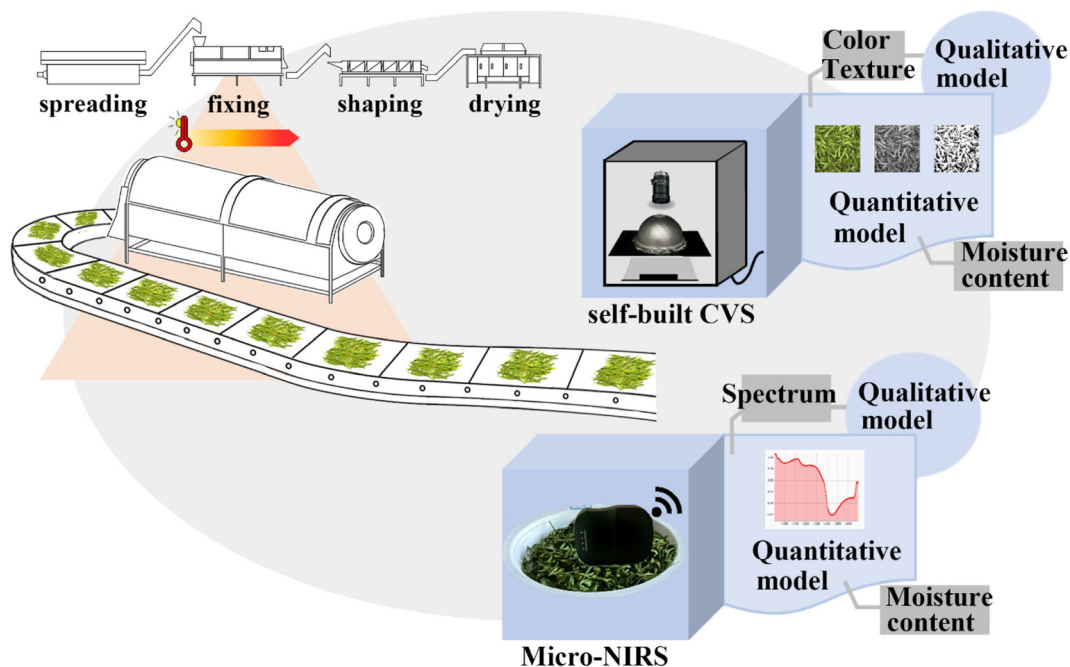


Figure 1. Green tea fixed sample collection process.

prevent interference from background or baseline drift during spectral acquisition.^{7,14} By comparing three preprocessing methods with the raw spectra, the optimal preprocessing method was selected.

Competitive adaptive reweighted sampling (CARS) is an effective characteristic wavelength optimization algorithm, and its detailed steps and principles are available in Jiang *et al.*¹⁵ In the present study, CARS was used to select the best combination of key wavelengths from the full spectrum as characteristic variables.

Determination of moisture content

The moisture content was determined based on the constant weight method at $103 \pm 2^\circ\text{C}$ according to the National Standard method (GB/T 8304-2013).

Chemometrics analysis

Division of collection samples

Prior to modeling, we divided the 120 fixed tea samples collected. By calculating the Euclidean distance between the samples using the Kennard–Stone method, the 120 samples were randomly divided into a calibration set (80) and a prediction set (40) with a ratio of 2:1. Eighty fixed tea samples of calibration set were used for modelling, and 40 independent fixed tea samples of prediction set were used for model validation.

Qualitative model establishment

Linear partial least squares discriminant analysis (PLS-DA), non-linear support vector machine (SVM) and non-linear least squares support vector machine (LS-SVM) models were established to discriminate the fixation degree. PLS-DA combines the PLS algorithm with discriminant analysis (DA) used to solve classification problems.¹⁶ In the PLS-DA model, the categorical variable Y is a vector holding category information. Samples were correctly classified when the predicted value differed from the actual value within ± 0.5 .¹⁷ SVM is a learning algorithm that maps the actual problem from a low-dimensional feature space to a high-dimensional

feature space through a kernel function. In the present study, a grid search optimization algorithm is used to optimize the two parameters, penalty factor c and the kernel function parameter g , in the range from 2^{-10} to 2^{10} to obtain accurate classification results.¹⁸ The LS-SVM algorithm is an optimized version of the SVM. LS-SVM uses a two-norm with equation constraints instead of the inequality constraints in the SVM algorithm in terms of the optimization algorithm.¹⁹ In the present study, the radial basis function (RBF), a commonly used kernel functions in LS-SVM, was used for modeling. A grid search was performed to optimize and determine two parameters, γ and σ^2 .²⁰

To measure the classification performance of the model, the correct classification rate (CCR), which is the ratio of the number of correctly classified samples to the total number of samples in the training or prediction set, is used for evaluation. The CCR takes values between 0% and 100%; the higher the CCR value in the calibration and prediction sets, the better the model performance.

Quantitative model establishment

The moisture content of samples were predicted using partial least squares regression (PLSR) and support vector regression (SVR). PLSR integrates the strengths of principal component analysis, typical correlation analysis and linear regression analysis. It projects high-dimensional data onto several independent latent variables (LVs) to solve the multicollinearity problem. The optimal LVs are determined by the minimum average predicted residual error sum of squares over five cross-validations.²¹ In PLS regression, regression coefficients (RCs) are calculated from the PLS loadings corresponding to the model with the optimal number of LVs and were used to assess the significance of each variable.²² SVR is an essential application branch of the SVM algorithm, which is better at minimizing structural risk during machine learning. RBF is commonly used as the SVR kernel function of choice because of its simple structure and fast operation speed; therefore, the RBF kernel function is used for model construction in the present study.¹⁸

The correlation coefficient (R_c) and root mean square error of cross-validation (RMSEC) of the calibration set and the correlation coefficient (R_p) and the root mean standard error of prediction (RMSEP) of the prediction set and the residual predictive deviation (RPD) were calculated to evaluate the performance of the respective models. The R indicates the degree of correlation between predicted and measured values, and RMSE indicates the degree of deviation between predicted and measured values.²³ Higher predictive model performance is shown by higher R_c and R_p and lower RMSEC and RMSEP.²⁴ The RPD is a comprehensive evaluation of model performance, the model delivers accurate predictions if $RPD \geq 2$.²⁵

Statistical analysis

MATLAB 2014a was used for data preprocessing, variable selection and modeling in Windows 10 (Microsoft Corp., Redmond, WA, USA). The statistical analysis of data was performed using SPSS, version 25.0 (IBM Corp, Armonk, NY, USA). The correlation analysis and plotting were performed using Origin 2021b (Origin Lab. Corp., Northampton, MA, USA).

RESULTS AND DISCUSSION

Definition of fixation degree of green tea

The sensory properties of green tea at different fixation temperatures were scored by professional tea reviewers (see Supporting information, Table S1). The obtained results revealed that the fixation temperature had a great impact on the quality of green tea, with varying degrees of variability in tea infusion color, aroma and taste. The highest scores for tea infusion color and aroma were achieved at 300 °C, and the best taste was achieved at 270 °C. However, there was no significant difference in the scores for each evaluation factor between the green tea samples at these two fixation temperature treatments. Compared to green tea samples under other temperature treatments, the fixation temperatures of 270 °C and 300 °C had higher scores and were therefore identified as the treatment temperatures that could produce moderately fixed green tea. Similar results were obtained in a study by Tan *et al.*²⁶ Moreover, the results of the sensory review team were consistent with the empirical evaluation by the tea makers on-site at the processing site.

Moisture content is a major factor in the formation of the appearance quality of tea.³ At a lower fixation temperature, the fresh leaves will contain too much moisture, resulting in more 'red stems', 'red leaves' and broken leaves in the finished tea, thus affecting the evaluation score. At a higher fixation temperature, the water in fresh leaves will rapidly turn into steam and expand, leaving blisters and scorch spots on the leaf surface, affecting the appearance and aroma of the tea. The moisture content of the samples was determined as shown in the Supporting information (Table S2). With increasing fixation temperature and deepening fixation, the moisture content significantly decreased. At fixation temperatures of 210–240 °C, the moisture content was high; at 330–360 °C, it was low; at 270–300 °C, it remained at a moderate moisture content of 55.1 g per 100 g.

In summary, samples with fixation temperatures of 210 °C and 240 °C, 270 °C and 300 °C, and 330 °C and 360 °C were identified as comprising insufficient, moderate and excessive fixation, respectively.

Image feature analysis

Figure 2(a) shows the original images of the green tea samples at different fixation temperatures. As the fixation process progressed, the leaf surface color changed from bright green to dark

green and grayish and lost luster, whereas the leaf texture changed from spreading to curling. This phenomenon is usually caused by the loss of water content in the leaves and changes in chlorophyll composition. To provide better identification of differences and more objective evaluation of image features, the color and texture change processes captured by the camera lens were quantified using image processing algorithms. The variation in color variable spaces and texture features during fixation is illustrated in Fig. 2(b–p). A similar downward trend in the values of the color variables R , G , B , H , V , L^* and b^* was observed as the fixation degree progressed, reflected in the color shift from a highly bright yellow–green to dark green. The S value sharply decreased at a moderate degree of fixation and rallied thereafter (Fig. 2h). By contrast, the a^* value sharply increased with increasing degree of processing, which is a result of the chlorophyll undergoing the magnesium removal reaction under the effect of high temperature.²⁷ The rate of magnesium removal increases with increasing temperature, and the brownish pheophytin makes the leaf color tend to become yellowish dark.²⁸ Observation of the texture feature values revealed a declining trend in the texture variables during the fixation process, except for the values of μ and U . The μ value responds to the degree of skewness of the image histogram, and the U value represents the local smoothness of the visible texture, with a significant increase in the value of both variables at moderate fixation.²²

Spectral information analysis

The raw spectral curves of 120 fixed leaf samples collected using micro-NIRS in the 900–1700 nm range are shown in Fig. 3(a). Figure 3(b, c) shows the average spectral curves for samples with different fixation temperatures and for samples with different fixation degrees; the abovementioned figures clearly show that the key to distinguishing the different treated samples is the intensity of the absorption values. The differences in absorption value intensities may be caused by discrepancies in the internal composition contents among the different treated samples.²⁹ The main absorption peaks of the spectral curves are found in the four wavelength ranges of 960–980 nm, 1160–1180 nm, 1430–1440 nm and 1650–1660 nm, as shown in Fig. 3(a–c). The absorption peak in the 960–980 nm range is regarded as a typical absorption peak for water molecules and is associated with the secondary-order frequency doubling of the O–H group in the water molecule.³⁰ The absorption peak at approximately 1180 nm is probably caused by the C–H secondary-order frequency doubling of CH₃ and the O–H first-order frequency doubling of the water molecule.³⁰ At approximately 1440 nm, the absorption peak is attributed to the first-order frequency doubling of O–H stretching.³¹ A peak at approximately 1650–1660 nm corresponds to the C–H secondary-order frequency doubling and S–H secondary-order frequency doubling of caffeine.³² In addition, it can be observed from the average spectral curves that the absorption values of the average spectral curves at 210–240, 270–300 and 330–360 °C tend to be of similar intensity, which also confirms the justification of our classification of the degree of fixation from the side.

Figure 3(d–f) shows the spectral curves after SNV, 1D and 2D preprocessing. After SNV preprocessing (Fig. 3d), the noise interference of the curve was weakened and the absorbance characteristics of the spectrum were strengthened. The overlapping peaks of the reshaped spectral curve after the derivative correction were separated (Fig. 3e, f), the resolution and sensitivity of the spectral signal were improved.³³

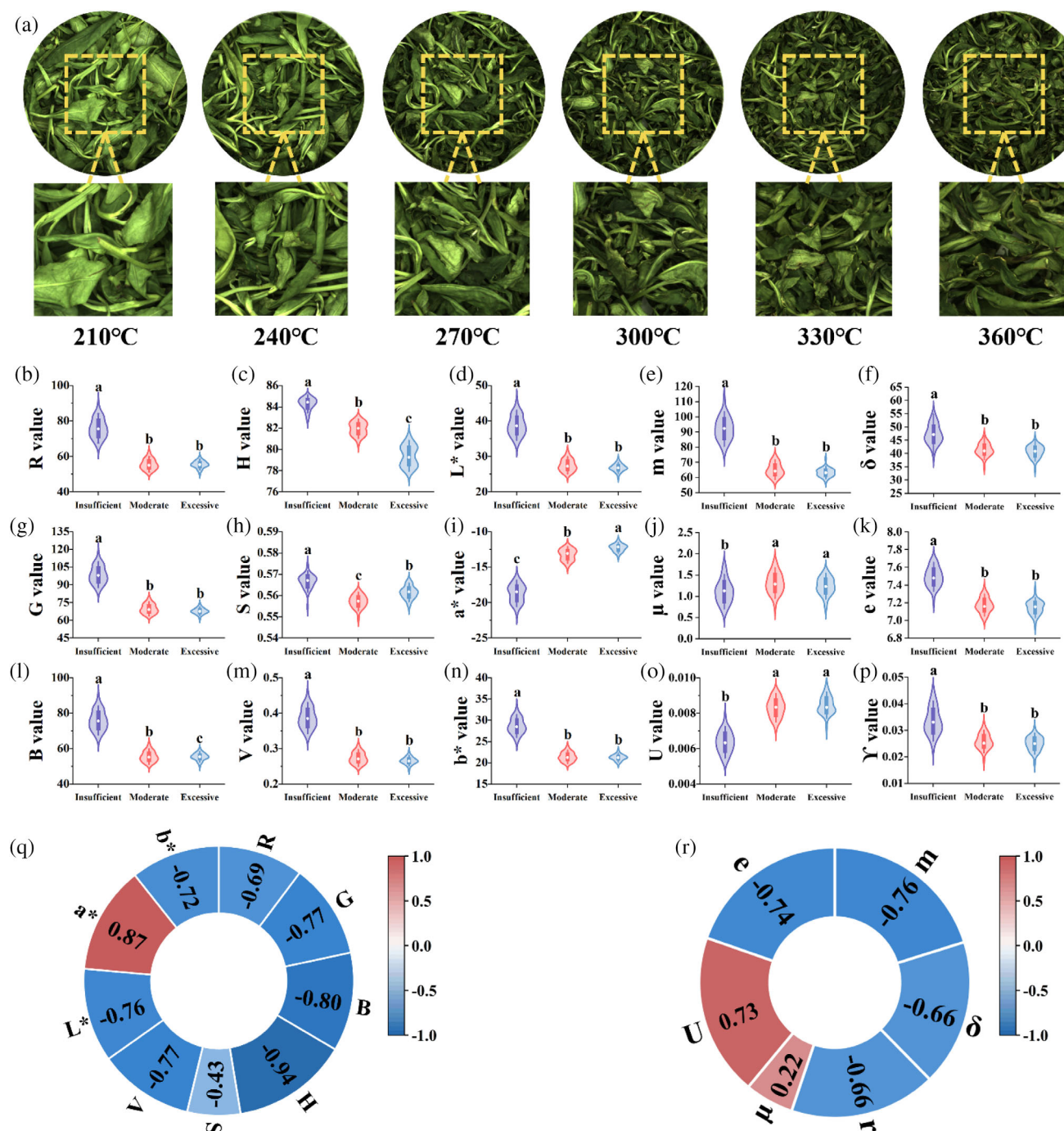


Figure 2. (a) Original images of green tea samples acquired by the computer vision system at different fixation temperatures and box-violin plots for each feature variable: (b) R ; (c) H ; (d) L^* ; (e) m ; (f) δ ; (g) G ; (h) S ; (i) a^* ; (j) μ ; (k) e ; (l) B ; (m) V ; (n) b^* ; (o) U ; (p) γ ; Spearman's two-tailed tests between the fixation degree and (q) color variables; (r) texture variables. Different lowercase letters indicate significant differences by Duncan's test ($P < 0.05$).

Green tea fixation discriminant models based on different sensors

The correlation coefficients between the image feature variables and the fixation degrees were obtained by Spearman's two-tailed test, as shown in Fig. 2(q, r). Except for the μ ($P < 0.01$, $r < 0.3$), which was not significantly correlated with the change in the degree of fixation, the a^* and the U were significantly positively correlated with the degree of fixation ($P < 0.01$, $r > 0.7$), whereas the remaining variables were negatively correlated. Eventually, R , G , B , H , V , L^* , a^* , b^* , m , δ , γ , U and e , which were significantly

correlated with the degree of fixation, were selected as the variables with image features ($P < 0.01$, $|r| > 0.6$). PLS-DA, SVM and LS-SVM classification algorithms were used to establish degree discrimination models based on texture, color and fusion feature (color and texture) variables of the fixed tea samples, respectively. As shown in Fig. 4(a), by comparing the performance of the discriminative models established with different algorithms, the non-linear LS-SVM model was shown to have the optimal discriminative ability in all three datasets of texture, color and fusion features, with the CCRs of prediction sets being 95.00%, 97.50% and

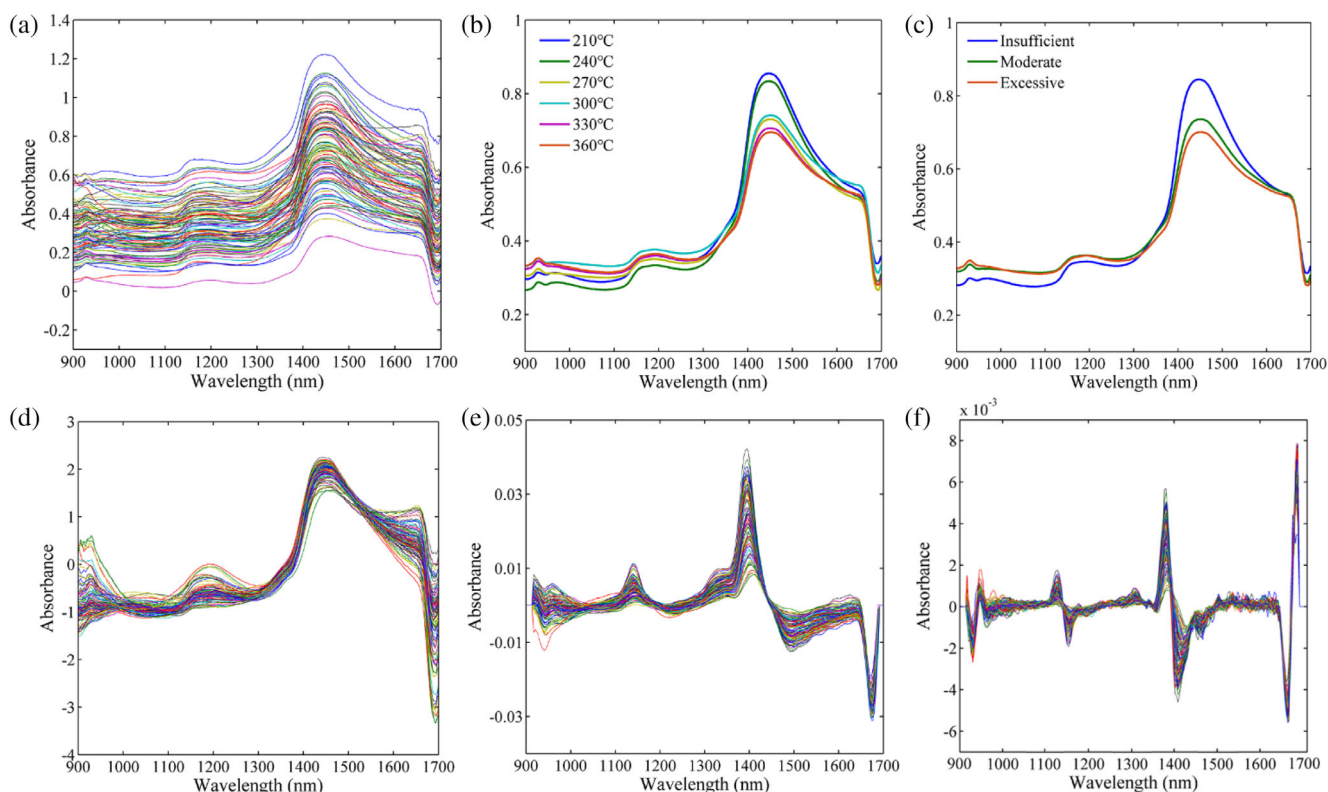


Figure 3. Spectrum of all tea samples: (a) raw spectrum; (b) average spectrum with different fixation temperatures; (c) average spectrum with different fixation degrees; (d) SNV preprocessed; (e) 1D preprocessed; and (f) 2D preprocessed.

100%, respectively. Of these, the most satisfactory performance occurred in the fusion features.

To increase the accuracy of the model, preprocessing and variable selection were performed on the raw data before building a discriminant model based on the spectral information. From the results shown in the Supporting information (Table S3), the three preprocessing methods all resulted in varying degrees of reduction in model accuracy. This finding indicates that the pre-processing methods may have eliminated potentially critical information at the same time as eliminating background interference. The raw spectral data were eventually used for subsequent analysis. The findings of the discriminant models based on the

17 characteristic wavelengths (see Supporting information, Table S4) optimally selected by CARS are shown in Fig. 4(b). The optimized spectral variables had varied effects on the prediction performance of the models: for the prediction samples, the linear PLS-DA model had a significant enhancement in accuracy, with a 17.5% point jump from the original CCR of 52.50%, although the results were still unsatisfactory. The non-linear SVM and LS-SVM models improved insignificantly but performed well overall, with CCRs of 82.50% and 100%, respectively. It is observed that the selection of characteristic wavelengths can effectively reduce the dimensionality of the data and remove the interference of irrelevant variables. The three classification models all demonstrated

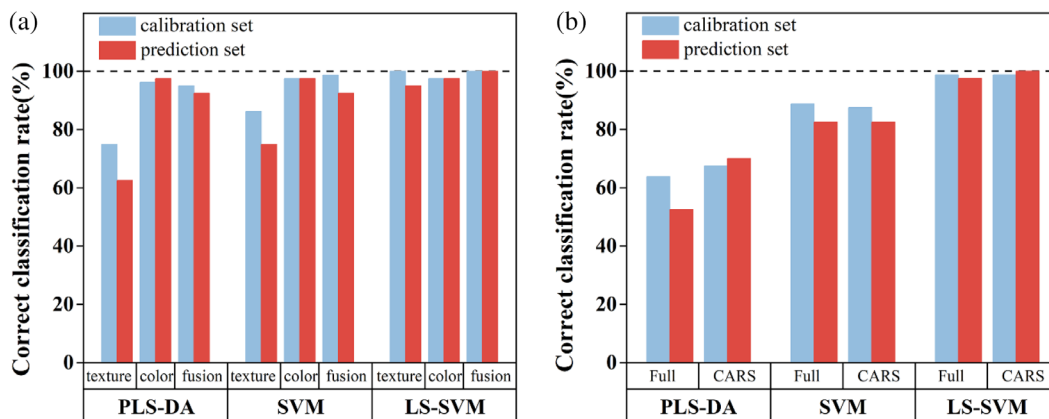


Figure 4. (a) Results of different fixation degree discrimination models based on single and fused image feature data. (b) Optimization results of the fixation degree model based on NIR spectral information.

feasibility in discriminating the degree of green tea fixation. However, the non-linear LS-SVM model achieved the lowest misclassification rate, with only one test sample misclassified in the calibration set (CCR = 98.75%) and 100% correct in the prediction set.

Moisture content prediction model analysis and comparison

Water is colorless, but differences in moisture content can affect the absorption and scattering of light by tea samples, ultimately resulting in different color contours. To clarify the degree of association between each image feature variable and the moisture content, Pearson correlation analysis was used for the analysis (Fig. 5a, b). The a^* , μ and U were negatively correlated with moisture content, while the rest of the feature variables were positively correlated. Excluding S and μ , which were weakly correlated with moisture content ($P < 0.01$, $|r| < 0.6$), similarly, R , G , B , H , V , L^* , a^* , b^* , m , δ , γ , U and e , which were significantly correlated with moisture content, were screened as image feature variables ($P < 0.01$, $|r| > 0.6$). Moreover, RCs of the PLSR models were used to evaluate the significance of each variable participating in modeling, as shown in the Supporting information (Fig. S1a). For moisture, the variables S and μ were not considered because of their lower absolute RC values.²³ This result is supported by the correlation analysis. The moisture content prediction models of the samples based on image feature variables were established using linear PLSR and non-linear SVR, and the results are shown in Table 1. The PLSR and SVR models based on texture feature variables were inferior in accuracy (R and RPD were lowest and RMSE was highest); the color feature variables obtained relatively satisfactory prediction accuracy in the modeling, especially for the SVR model ($R_c = 0.9903$, RMSEC = 0.0115, $R_p = 0.9877$, RMSEP = 0.0111, RPD = 6.17); the fused feature variables achieved the optimal moisture content prediction results in the SVR model ($R_c = 0.9927$, RMSEC = 0.0099, $R_p = 0.9884$, RMSEP = 0.0106, RPD = 6.46), effectively improving model accuracy and model stability.

The results of the original and the preprocessed spectra of the SVR moisture prediction model were compared (see Supporting information, Table S5). The second derivative (2D) preprocessing improved the accuracy and overall generalization capability of the model in prediction sets (lower RMSEP, higher R_p) and the RPD improved from the original 1.80 to 2.20. Subsequently, the

moisture prediction models were optimized, as shown in Table 2. CARS selected 21 valid wavelengths associated with moisture content (see Supporting information, Table S6). Combined with the regression coefficient plot of the PLSR model (see Supporting information, Fig. S1b), the absorption peaks with higher RC values at approximately 950 nm and 1440 nm, respectively, were mainly related to the stretching vibration of the O—H group in the water molecule, which had some overlap with valid wavelengths. Based on the valid wavelengths, simpler prediction models were established, at the same time as effectively improving the prediction accuracy of all models during model validation. The linear PLSR model achieved the best prediction accuracy at $R_c = 0.8880$, RMSEC = 0.0388, $R_p = 0.9369$, RMSEP = 0.0265 and RPD = 2.59, and the non-linear SVR model achieved the best prediction accuracy at $R_c = 0.9347$, RMSEC = 0.0300, $R_p = 0.9427$, RMSEP = 0.0221 and RPD = 3.10. Clearly, the SVR model exhibited a greater advantage over the PLSR model as a result of the non-linear relationship between the moisture content and spectral information during fixation.

The feasibility validation results of the optimal moisture content prediction models obtained based on two sensors are shown in the Supporting information (Table S7). The independent samples t -test result showed no statistical difference between the measured and predicted values ($P > 0.05$).

Sensor comparison and discussion

Moisture is an essential parameter in the fixation process of green tea. The level of moisture content in samples with different fixation degrees directly affects the quality of fixed tea samples, most intuitively reflected in the differences in sensory properties. The present study used spectral information acquired by micro-NIRS and image information acquired by CVS, combined with chemometric algorithms, to intelligently monitor the fixation quality of green tea. By comparing the results of different discriminant models, the non-linear CARS-LS-SVM model based on spectral information achieved an accurate judgment of the fixation degree (CCR of 100% for the prediction set). This result was consistent with the study by Jin *et al.*,³⁴ who demonstrated that a non-linear model (SVM) achieved higher performance than linear models (K-nearest neighbor and linear discriminant analysis) with respect to discriminating the quality of black tea in-process products. Likewise, the advantage of the non-linear model was shown in image

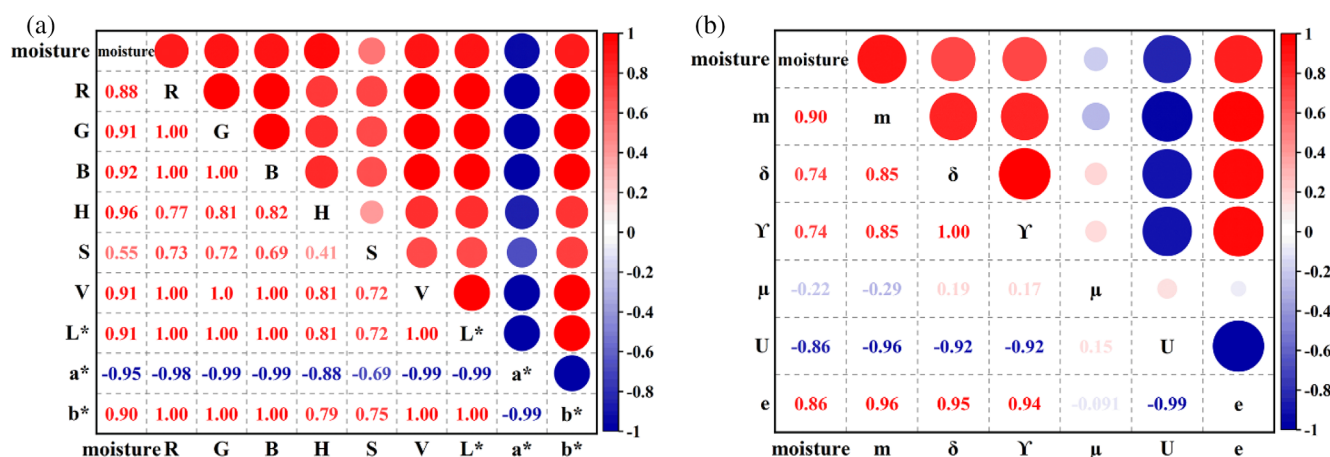


Figure 5. Pearson correlation analysis between moisture content and (a) color variables and (b) texture variables.

Table 1. Performance of PLSR and SVR models based on single and fused image feature data for predicting moisture content

Component	Model	Data	Feature numbers	Parameters	Calibration set		Prediction set			Slope	Bias
					<i>R</i> _c	RMSEC	<i>R</i> _p	RMSEP	RPD		
Moisture	PLSR	texture features	6	LVs = 3	0.9196	0.0306	0.8993	0.0327	2.09	0.83	0.10
		color features	9	LVs = 5	0.9860	0.0134	0.9818	0.0139	4.93	0.97	0.01
		data fusion	13	LVs = 5	0.9866	0.0133	0.9880	0.0112	6.12	0.98	0.01
	SVR	texture features	6	<i>c</i> = 0.33, <i>g</i> = 256	0.9462	0.0254	0.9150	0.0305	2.25	0.85	0.09
		color features	9	<i>c</i> = 0.57, <i>g</i> = 256	0.9903	0.0115	0.9877	0.0111	6.17	0.98	0.01
		data fusion	13	<i>c</i> = 1, <i>g</i> = 48.50	0.9927	0.0099	0.9884	0.0106	6.46	0.98	0.01

Table 2. Optimization results of PLSR and SVR models based on NIR spectral information for predicting moisture content

Component	Model	Preprocessing	Variable selection methods	Variable numbers	Parameters	Calibration set		Prediction set			Slope	Bias
						<i>R</i> _c	RMSEC	<i>R</i> _p	RMSEP	RPD		
Moisture	PLSR	2-Derivative	Full	228	LVs = 4	0.7998	0.0499	0.8522	0.0381	1.80	0.74	0.14
			CARS	21	LVs = 5	0.8880	0.0388	0.9369	0.0265	2.59	0.87	0.07
	SVR	2-Derivative	Full	228	<i>c</i> = 6.96, <i>g</i> = 1024	0.9434	0.0287	0.8944	0.0311	2.20	0.82	0.10
			CARS	21	<i>c</i> = 64, <i>g</i> = 1024	0.9347	0.0300	0.9427	0.0221	3.10	0.91	0.05

information, where the LS-SVM model based on fusion features obtained the same discrimination accuracy for fixation degree as the spectral information (CCR = 100%). These results showed that the variations in the information on the appearance and internal composition of tea at different fixation stages are dynamic and non-linear, and therefore, accurate discrimination is difficult to realize with a linear algorithm (PLS-DA).

In predicting the moisture content of fixed tea samples, the proposed 2D-CARS-SVR model based on spectral information exhibited significant superiority in achieving accurate prediction of moisture content (*R*_c = 0.9347, RMSEC = 0.0300, *R*_p = 0.9427, RMSEP = 0.0221, RPD = 3.10). Unlike the moisture prediction model based on the whole green tea process by Liu *et al.*,³⁵ the present study targeted the fixation process. We established a model with a wider range of moisture content by collecting a larger sample size. Our prediction accuracy was higher with the RMSEP of 0.0221 compared to their model RMSEP of 0.0460. Observing the results of image information modeling, the SVR model, established by fusion feature information, provided the optimal moisture content prediction results with *R*_c = 0.9927, RMSEC = 0.0099, *R*_p = 0.9884, RMSEP = 0.0106 and RPD = 6.46. Clearly, the multidimensional fusion features provide a more comprehensive description of the variations in the green tea fixation samples. The superiority of fusion features has also been demonstrated in a quantitative evaluation of black tea withering quality.¹⁰

Compared to the spectral information, the image information performs better in the process monitoring of the fixation quality, particularly in moisture prediction capability. The reason for this result may be that the change in moisture content during fixation is most intuitively reflected in the altered appearance quality, and NIRS provides point-based spectral information, lacking more comprehensive spatial information;³⁶ in contrast, CVS-captured images are more advantageous. Additionally, NIRS is susceptible to heat and humidity process environments and sample homogeneity states.⁶ This conclusion was also verified in the following studies, where the moisture prediction results of An *et al.*,³⁷ using CVS for withering process samples (*R*_p = 0.9957, RMSEP = 0.0059)

were better than those of Shen *et al.*⁷ using micro-NIRS (*R*_p = 0.9931, RMSEP = 0.0095).

Here, we first validated and compared the performance of two cost-effective sensors, micro-NIRS and CVS, for the fixation quality monitoring of green tea, illustrating their suitability as intelligent monitoring tools and their potential for industrial applications from multiple perspectives. However, the present study has limitations because we selected only one green tea suitable variety for single-batch experiments. Next, factors such as variety and batch will be taken into account to expand the sample database and improve the generalization ability of the model. In summary, the present study provides optional intelligent monitoring means for standardized production in the green tea industry.

CONCLUSIONS

The present study explores the feasibility of low-cost micro-NIRS and CVS for the rapid at-line monitoring of green tea fixation quality, comparing the performance both qualitatively and quantitatively of the models constructed from spectral and image information. Both the spectral-based and the image-based LS-SVM models achieved 100% discrimination accuracy in fixation discrimination. However, in moisture content prediction, the image-based SVR model provided better prediction results than the spectral information. Additionally, image feature fusion and spectral variable selection methods further optimize the accuracy of qualitative/quantitative models. The results of the present study provide theoretical support and technical guidance for the automation of the green tea fixation process, which can be applied in the future for real-time monitoring of green tea processing lines.

ACKNOWLEDGEMENTS

This study was supported by the National Key Research and Development Program of China (2021YFD1601100).

CONFLICTS OF INTEREST

The authors declare that they have no conflicts of interest.

SUPPORTING INFORMATION

Supporting information may be found in the online version of this article.

REFERENCES

- Li HX, Zhong Q, Luo FJ, Wang XR, Zhou L, Chen ZM *et al.*, Residue degradation and metabolism of spinetoram in tea: a growing, processing and brewing risk assessment. *Food Control* **125**:107955 (2021a).
- Li J, Hua JJ, Yuan HB, Deng YL, Zhou QH, Yang YQ *et al.*, Investigation on green tea lipids and their metabolic variations during manufacturing by nontargeted lipidomics. *Food Chem* **339**:128114 (2021b).
- Panigrahi N, Bhol CS and Das BS, Rapid assessment of black tea quality using diffuse reflectance spectroscopy. *J Food Eng* **190**:101–108 (2016).
- Wei YZ, Wu FY, Xu J, Sha JJ, Zhao ZF, He Y *et al.*, Visual detection of the moisture content of tea leaves with hyperspectral imaging technology. *J Food Eng* **248**:89–96 (2019).
- Chen QS, Zhang DL, Pan WX, Ouyang Q, Li HH, Urmila K *et al.*, Recent developments of green analytical techniques in analysis of tea's quality and nutrition. *Trends Food Sci Technol* **43**:63–82 (2015).
- Porep JU, Kammerer DR and Carle R, On-line application of near infrared (NIR) spectroscopy in food production. *Trends Food Sci Technol* **46**:211–230 (2015).
- Shen S, Hua JJ, Zhu HK, Yang YQ, Deng YL, Li J *et al.*, Rapid and real-time detection of moisture in black tea during withering using micro-near-infrared spectroscopy. *LWT* **155**:112970 (2022).
- Jin G, Wang YJ, Li MH, Li TH, Huang WJ, Li LQ *et al.*, Rapid and real-time detection of black tea fermentation quality by using an inexpensive data fusion system. *Food Chem* **358**:129815 (2021).
- Dong CW, Liang GZ, Hu B, Yuan HB, Jiang YW, Zhu HK *et al.*, Prediction of Congou black tea fermentation quality indices from color features using non-linear regression methods. *Sci Rep* **8**:10535 (2018a).
- Liang GZ, Dong CW, Hu B, Zhu HK, Yuan HB, Jiang YW *et al.*, Prediction of moisture content for Congou black tea withering leaves using image features and nonlinear method. *Sci Rep* **8**:7854 (2018).
- Chen C, Liu B, Song FH, Jiang JJ, Li ZF, Song CF *et al.*, An adaptive fuzzy logic control of green tea fixation process based on image processing technology. *Biosyst Eng* **215**:1–20 (2022).
- Liu D, Pu HB, Sun DW, Wang L and Zeng XA, Combination of spectra and texture data of hyperspectral imaging for prediction of pH in salted meat. *Food Chem* **160**:330–337 (2014).
- Dong CW, Li J, Wang JJ, Liang GZ, Jiang YW, Yuan HB *et al.*, Rapid determination by near infrared spectroscopy of theaflavins-to-thearubigins ratio during Congou black tea fermentation process. *Spectrochim Acta, Part A* **205**:227–234 (2018b).
- Ren GX, Liu Y, Ning JM and Zhang ZZ, Hyperspectral imaging for discrimination of Keemun black tea quality categories: multivariate calibration analysis and data fusion. *Int J Food Sci Technol* **56**:2580–2587 (2021a).
- Jiang H, Xu WD, Ding YH and Chen QS, Quantitative analysis of yeast fermentation process using Raman spectroscopy: comparison of CARS and VCPA for variable selection. *Spectrochim Acta, Part A* **228**:117781 (2020).
- Gowen AA, Downey G, Esquerre C and O'Donnell CP, Preventing overfitting in PLS calibration models of near-infrared (NIR) spectroscopy data using regression coefficients. *J Chemom* **25**:375–381 (2011).
- Wang YJ, Li MH, Li LQ, Ning JM and Zhang ZZ, Green analytical assay for the quality assessment of tea by using pocket-sized NIR spectrometer. *Food Chem* **345**:128816 (2021a).
- Chang CC and Lin CJ, LIBSVM: a library for support vector machines. *Acm Trans Intell Syst Technol* **27**:1–27 (2011).
- Gestel TV, Suykens JAK, Baesens B, Viaene S, Vanthienen J, Dedene G *et al.*, Benchmarking least squares support vector machine classifiers. *Mach Learn* **54**:5–32 (2004).
- Wang YJ, Li TH, Jin G, Wei YM, Li LQ, Kalkhajah YK *et al.*, Qualitative and quantitative diagnosis of nitrogen nutrition of tea plants under field condition using hyperspectral imaging coupled with chemometrics. *J Sci Food Agric* **100**:161–167 (2020).
- Wang YJ, Li LQ, Liu Y, Cui QQ, Ning JM and Zhang ZZ, Enhanced quality monitoring during black tea processing by the fusion of NIRS and computer vision. *J Food Eng* **304**:110599 (2021b).
- Li LQ, Xie SM, Ning JM, Chen QS and Zhang ZZ, Evaluating green tea quality based on multisensor data fusion combining hyperspectral imaging and olfactory visualization systems. *J Sci Food Agric* **99**:1787–1794 (2019).
- Wang YJ, Ren ZY, Chen YY, Lu YC, Deng WW, Zhang ZZ *et al.*, Visualizing chemical indicators: spatial and temporal quality formation and distribution during black tea fermentation. *Food Chem* **401**:134090 (2023).
- Ouyang Q, Zhao JW, Pan WX and Chen QS, Real-time monitoring of process parameters in rice wine fermentation by a portable spectral analytical system combined with multivariate analysis. *Food Chem* **190**:135–141 (2016).
- Chang CW, Laird DA, Mausbach MJ and Hurburgh CR, Near-infrared reflectance spectroscopy–principal components regression analyses of soil properties. *Soil Sci Soc Am J* **65**:480–490 (2001).
- Tan JF, Jin HQ, Huang YJ, Peng QH, Shao Q, Shen X *et al.*, Design and application of automatic production line on roasted green tea. *J Tea Sci* **30**:229–234 (2021) (in Chinese).
- Lanfer-Marquez UM, Barros RMC and Sinnecker P, Antioxidant activity of chlorophylls and their derivatives. *Food Res Int* **38**:885–891 (2005).
- Wu SM, Jiang YW, Hua JJ, Yuan HB and Li J, Effects of different full drying methods on the chlorophyll degradation in green tea as analyzed by UPLC-Q-Exactive Orbitrap-MS. *Food Sci* **43**:44–51 (2022) (in Chinese).
- Ren GX, Wang YJ, Ning JM and Zhang ZZ, Highly identification of keemun black tea rank based on cognitive spectroscopy: near infrared spectroscopy combined with feature variable selection. *Spectrochim Acta, Part A* **230**:118079 (2020).
- Yi QX, Bao AM, Wang Q and Zhao J, Estimation of leaf water content in cotton by means of hyperspectral indices. *Comput Electron Agric* **90**:144–151 (2013).
- Lee MS, Hwang YS, Lee J and Choung MG, The characterization of caffeine and nine individual catechins in the leaves of green tea (*Camellia sinensis* L.) by near-infrared reflectance spectroscopy. *Food Chem* **158**:351–357 (2014).
- Chen QS, Chen M, Liu Y, Wu JZ, Wang XY, Ouyang Q *et al.*, Application of FT-NIR spectroscopy for simultaneous estimation of taste quality and taste-related compounds content of black tea. *J Food Sci Technol* **55**:4363–4368 (2018).
- Ren GX, Liu Y, Ning JM and Zhang ZZ, Assessing black tea quality based on visible-near infrared spectra and kernel-based methods. *J Food Compos Anal* **98**:103810 (2021b).
- Jin G, Wang YJ, Li LQ, Shen SS, Deng WW, Zhang ZZ *et al.*, Intelligent evaluation of black tea fermentation degree by FT-NIR and computer vision based on data fusion strategy. *LWT* **125**:109216 (2020).
- Liu ZY, Zhang RT, Yang CS, Hu B, Luo X, Li Y *et al.*, Research on moisture content detection method during green tea processing based on machine vision and near-infrared spectroscopy technology. *Spectrochim Acta, Part A* **271**:120921 (2022).
- Lohumi S, Lee S, Lee H and Cho BK, A review of vibrational spectroscopic techniques for the detection of food authenticity and adulteration. *Trends Food Sci Technol* **46**:85–98 (2015).
- An T, Yu H, Yang CS, Liang GZ, Chen JY, Hu ZH *et al.*, Black tea withering moisture detection method based on convolution neural network confidence. *J Food Process Eng* **43**:e13428 (2020).