

Research Article

Near Infrared Spectroscopy Based on Supervised Pattern Recognition Methods for Rapid Identification of Adulterated Edible Gelatin

Hao Zhang ¹, Haifeng Sun,¹ Ling Wang ¹, Shun Wang,¹ Wei Zhang,¹ and Jiandong Hu^{1,2}

¹College of Mechanical and Electrical Engineering, Henan Agricultural University, Zhengzhou 450002, China

²State Key Laboratory of Wheat and Maize Crop Science, Zhengzhou 45002, China

Correspondence should be addressed to Hao Zhang; hao.zhang2016@hotmail.com

Received 9 October 2018; Revised 10 November 2018; Accepted 4 December 2018; Published 20 December 2018

Academic Editor: Arnaud Cuisset

Copyright © 2018 Hao Zhang et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

The aim of this work is to identify the adulteration of edible gelatin using near-infrared (NIR) spectroscopy combined with supervised pattern recognition methods. The spectral data obtained from a total of 144 samples consisting of six kinds of adulterated gelatin gels with different mixture ratios were processed with multiplicative scatter correction (MSC), Savitzky–Golay (SG) smoothing, and min-max normalization. Principal component analysis (PCA) was first carried out for spectral analysis, while the six gelatin categories could not be clearly distinguished. Further, linear discriminant analysis (LDA), soft independent modelling of class analogy (SIMCA), backpropagation neural network (BPNN), and support vector machine (SVM) were introduced to establish discrimination models for identifying the adulterated gelatin gels, which gave a total correct recognition rate of 97.44%, 100%, 97.44%, and 100%, respectively. For the SIMCA model with significant level $\alpha = 0.05$, sample overlapping clustering appeared; thus, the SVM model presents the best recognition ability among these four discrimination models for the classification of edible gelatin adulteration. The results demonstrate that NIR spectroscopy combined with unsupervised pattern recognition methods can quickly and accurately identify edible gelatin with different adulteration levels, providing a new possibility for the detection of industrial gelatin illegally added into food products.

1. Introduction

Edible gelatin is produced by hydrolyzed collagen extracted from fresh skin and bones of animals (such as pig, cattle, or fish) through a series of complicated procedures, which is rich in 18 kinds of amino acids needed by the human body. Due to its high protein, no fat, and no cholesterol, edible gelatin is widely used as a food thickener in the food industry (such as jelly, yogurt, and galantine) [1]. Industrial gelatin is commonly made from waste leather by decoloring, bleaching, and washing, which requires a large amount of chrome-containing tanning agents, so the heavy metal chromium content in industrial gelatin seriously exceeds the standard. Once the human body ingests a large amount of heavy metal chromium, it will cause kidney damage, and in serious cases, may lead to cancer [2]. Therefore, the Chinese

government prohibits industrial gelatin from being used in food and pharmaceutical products, and the national standards for “edible gelatin” clearly stipulate that the maximum allowable limit of the chromium content in edible gelatin is 2.0 mg/kg. However, due to the low price and simple production of industrial gelatin, some unscrupulous merchants added industrial gelatin to food products instead of edible gelatin [3]. In China, food safety accident is in high incidence, such as the “poison capsule” and “old yogurt” incidents reported in 2012, as well as the “poison bean jelly” incident reported in 2014. At present, there is no standard detection method for industrial gelatin illegally added in food products, so a fast and nondestructive method is needed to identify adulteration of edible gelatin.

Currently, some analytical methods for the identification of gelatin in products include the electrophoretic method

[4, 5], enzyme linked immune sorbent assay (ELISA) [6–8], high performance liquid chromatography (HPLC) [9–11], and polymerase chain reaction (PCR) [12–15]. However, most of these methods are expensive and time consuming although the sensitivity and accuracy are high. In recent years, Fourier transmission infrared spectroscopy (FTIRS) has been used as a nondestructive spectroscopic analytical method to differentiate different gelatin origins, for example, Hashim et al. [16] combined FTIR with attenuated total reflectance (ATR) to obtain a clear distinction between gelatin samples of bovine and porcine origins, and Cebi et al. [17] used FTIRS combined with principal component-cluster analysis to achieve a better distinction between pig gelatin, bovine gelatin, and fish gelatin.

However, most of these methods were used to differentiate gelatin origins by combining with chemometric methods; few studies were performed to explore if the edible gelatin was adulterated with the toxic industrial gelatin. As a fast and nondestructive method, NIR spectroscopy has been widely used in the detection of adulteration identification of agricultural products and foods [18–22]. In this paper, NIR spectroscopy combined with four kinds of supervised pattern recognition methods was used to study the feasibility of identifying edible gelatin adulterated with industrial gelatin. For this purpose, NIR spectra expressed as absorbance were measured on six kinds of gelatin gels with different adulteration ratios, principle component analysis (PCA) was first used for exploratory analysis, and then discrimination models based on linear discriminant analysis (LDA), soft independent modelling of class analogy (SIMCA), back-propagation neural network (BPNN), and support vector machine (SVM) methods were developed, where two parameters of recognition rate and rejection rate were used to assess the predictive performance of discrimination models.

2. Materials and Methods

2.1. Sample Preparation. The edible gelatin and industrial gelatin used in the experiment were purchased from Henan Boyang Biotechnology Co., Ltd (see Table S1). Each sample was weighed 10 g to prepare for gelatin gels with six different mixture ratios, i.e., 10 g edible gelatin, 8 g edible gelatin + 2 g industrial gelatin, 6 g edible gelatin + 4 g industrial gelatin, 4 g edible gelatin + 6 g industrial gelatin, 2 g edible gelatin + 8 g industrial gelatin, and 10 g industrial gelatin, which were denoted by pure edible gelatin (Class 1), 4:1 adulterated gelatin (Class 2), 3:2 adulterated gelatin (Class 3), 2:3 adulterated gelatin (Class 4), 1:4 adulterated gelatin (Class 5), and pure industrial gelatin (Class 6), respectively, as shown in Table S2 and Figure S1.

The preparation process of gelatin gels is as follows: (1) the mixed 10 g gelatin sample was placed in a 250 mL beaker and added with 100 mL purified water to soak for about half an hour; (2) the beaker was placed in a water bath pan for heating by setting the temperature at 50°C, and the solution was stirred while heating until the gelatin was completely dissolved; (3) after cooling to room temperature, the gelatin solution was poured into a small square mold and then placed in a refrigerator to cool about 5 hours for shaping.

2.2. Spectral Acquisition. The NIR measuring setup used in this experiment was constructed with a visible near-infrared halogen lamp light source, an optical fiber adjustable attenuator, two multimode optical fiber with core diameter of 600 μm , a transfection mount, an NIR spectrometer (NIR Quest 256–2.5, Ocean Optics, USA), and a computer. The wavelength range of the light source is 350–2400 nm, and the spectral range of the NIR spectrometer is 900–2500 nm. Data points were collected every 3 nm, with 512 data points for each spectral curve. In the experiment, the NIR absorption spectrum was measured in a transmission geometry and expressed as absorbance for spectral analysis. Each sample was measured at 5 different positions and its average spectrum was taken as the NIR spectrum of the sample.

2.3. Spectral Analysis. To establish a feasible NIRS identification model, some preprocessing methods, including multiplicative scatter correction (MSC), Savitzky–Golay (SG) smoothing, and min-max normalization, were applied to reduce the background noise and increase the signal-noise ratio. In this study, as one of the most commonly used unsupervised machine learning method, PCA was first employed for identification analysis, by reducing the high dimension of the original spectral matrix to a new low-dimensional spectral matrix with the minimal loss of information [23]. Afterwards, four kinds of supervised pattern recognition methods, LDA, SIMCA, BPNN, and SVM, were introduced to establish the discrimination models, respectively.

LDA is to find out a so-called discriminant function which best separates the classes by minimizing the distance of within-class samples and maximizing the distance of between-class samples [24]. SIMCA is based on establishing a PCA model for each class, where two critical values of the Euclidean distances and Mahalanobis distances assessed separately by the residual Q statistic and the Hotelling's T^2 statistic are used to determine boundaries around the samples belonging to one particular class [25, 26]. BPNN is an artificial neural network (ANN) based on the back-propagation method; to obtain the output of a well-trained network, the errors of each layer are calculated layer by layer from the beginning of the output layer and then used to optimize the network structure by adjusting the weights of each layer [27]. SVM achieves the goal of classification by using a kernel function to solve a linear or nonlinear problem, and the commonly used kernel functions include linear kernel, polynomial kernel, Gaussian kernel, and radial basis function; the generalization performance of SVM is highly related to the kernel function and the parameters in kernel function [28].

2.4. Discrimination Model Evaluation. In order to evaluate the performance of the established discrimination models, two parameters of recognition rate and rejection rate were proposed to reflect the reliability of clustering among different classes. The recognition rate refers to the ratio of the number of samples identified by one certain class to the total number of samples from the class, while the rejection rate

refers to the ratio of the number of samples rejected by one certain class to the total number of samples from other classes. When both values of recognition rate and rejection rate are 100%, there is no overlap between the different classes of samples, meaning that they could be fully distinguished.

2.5. Software. The preprocessing of NIR spectra was carried out by the self-designed MATLAB programs. LDA and SIMCA models were built using the Unscrambler X 10.4 software (Camo AS, Oslo, Norway), the BPNN model was built using MATLAB programs based on ANN toolbox, and SVM models were built using LibSVM toolbox developed by professor Lin Zhiren from Taiwan University [29]; all MATLAB programs were run under MATLAB R 2016A software (Mathworks Inc., Natick, MA, USA).

3. Results and Discussions

3.1. Spectral Characteristics of Gelatin Gels. After the preprocessing of MSC, SG smoothing, and min-max normalization, the spectra improved a lot, as shown in Figure 1, displaying the mean values and standard deviations for each group of gelatin, where the standard deviations of spectra are very small. As can be seen clearly, in the wavelength range of 800–2200 nm, the normalized spectra have a similar spectral shape but different intensities. In addition, gelatin spectra show some prominent characteristic absorption peaks. The peaks approximating to 940 nm, 1490 nm, and 1930 nm are mainly caused by the overtones and combination bands of O–H group in water molecules. The peaks at about 1200 nm and 1730 nm should be generated by the overtones and combination bands of C–H group in gelatin samples, while the multiple peaks at 2000–2200 nm are supposed to primarily stem from the combination bands of C–H and N–H [30–32]. Due to the similarity of the spectral shapes, it is very difficult to straightforwardly determine if edible gelatin was added with industrial gelatin only from NIR spectrum.

3.2. Discrimination Models. In order to establish an appropriate discrimination model and evaluate its validity, the original data were divided into a training set and a validation set, where the training set was used to establish discrimination models and the validation set was used to verify the validity of the models. To improve the prediction ability of the models, according to a ratio of 2:1, 75 samples were randomly selected from a total of 114 samples of different gelatin varieties as a training set, and the remaining 39 samples were utilized as validation set, as shown in Tables 1 and S2.

3.2.1. LDA Model. In this work, PCA was first performed on the normalized spectral data; the first three principal components PC1, PC2, and PC3 accounted for 93.24%, 5.62%, and 0.51% of the total variance, respectively, making a 99.37% of the cumulative total variance, which indicates that the first three PCs can represent most of the

NIR spectral information of gelatin gels. The loading plot and score plot of the first three PCs are shown in Figure 2. In the loading plot (Figure 2(a)), the spectral shape of the pure PC1 loading vector exhibits most of the characteristic absorption peaks shown in Figure 1, although both positive and negative contributions are included. PC2 and PC3 also give the positive and negative contributions, shifting the intensity maximum of the pure PC1 spectrum toward to the real spectrum. Figure 2(b) shows the score scatter plot of the first two and three PCs, respectively; it can be seen that some samples belonging to different classes are clustering, while others belonging to the same class are dispersing into two clusters, resulting in that the six classes of gelatin gels cannot be clearly distinguished due to the vague boundaries.

After principal component analysis, the optimal principal component number of 8 was determined by leave-one-out cross validation (LOOCV) and selected as the input variables of the LDA model, and the training set was selected to establish the LDA recognition model for gelatin adulteration. Then, the model was applied to the training set and the validation set, as shown in Table 2. The results show that both the recognition rate and rejection rate in the training set are 100%, while in the validation set, the recognition rate and rejection rate of other kinds of gelatin except pure edible gelatin and 4:1 adulterated gelatin are 100%, which means that these four kinds of gelatin can be well recognized. In the validation set, as shown in Table S2, a sample belonging to pure edible gelatin was incorrectly identified as 4:1 adulterated gelatin, resulting in a recognition rate of 90.91% for pure edible gelatin and a rejection rate of 96.88% for 4:1 adulterated gelatin. Since one sample in the validation set was wrongly classified, it led to a total correct recognition rate of 97.44%. The results show that the LDA discrimination model has a good classification ability on gelatin adulterated samples.

3.2.2. SIMCA Model. The SIMCA model was built based on the PCA model for each class of adulterated gelatin gels in the training set, and then the PCA models were individually performed on each class in the validation set for classification. For the PCA models, the optimal number of PCs was determined by the cross validation method (Table 3). Table 3 shows the predicting results of the SIMCA model in the training set and validation set when significance level $\alpha = 0.05$ and $\alpha = 0.1$, respectively.

As can be seen clearly, when the significance level is 0.05, except for the pure edible gelatin, the recognition rate and rejection rate for each class were 100% both in the training set and validation set. For the pure edible gelatin, the recognition rate was also 100%, while the rejection rate was 91% and 82%, separately for the training set and validation set. As listed in Tables S4 and S5, six samples in the training set belonging to 4:1 adulterated gelatin were incorrectly identified as pure edible gelatin, while five samples in the validation set were misclassified. When the significance level is 0.1, the recognition rate of 100% and rejection rate of 98% were obtained for the training set, indicating that one sample was overlapping

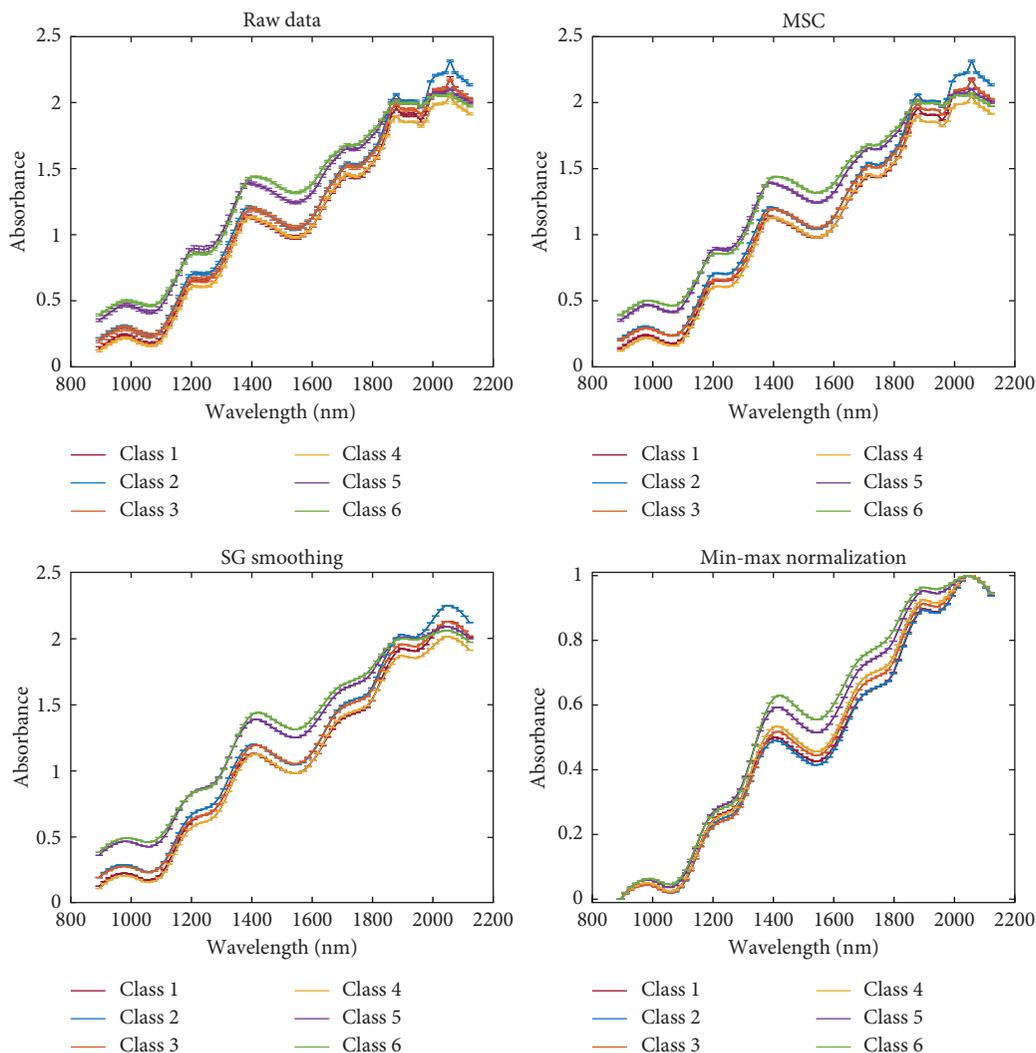


FIGURE 1: Mean values and standard deviations of NIR spectra for each group of gelatin before and after preprocessing.

TABLE 1: Training set and validation set of gelatin samples with six different adulteration levels.

Category	Training set	Validation set	Total number
Pure edible gelatin	9	11	20
4:1 adulterated gelatin	13	7	20
3:2 adulterated gelatin	13	5	18
2:3 adulterated gelatin	13	7	20
1:4 adulterated gelatin	12	6	18
Pure industrial gelatin	15	3	18
Total number	75	39	114

which appeared in two classes (Table S6). While in the validation set, the recognition rate and rejection rate for each class were 100% (Table S7), meaning that the SIMCA model can fully recognize the six classes of gelatin gels. Therefore, although 100% of the recognition rate can be achieved for the SIMCA model, the rejection rate less than 100% in particular for the prediction set indicates the possible presence of clustering overlap between the first two classes.

3.2.3. BPNN Model. Since the whole spectral band from 894 nm to 2302 nm consists of 432 data points, the number of input nodes is 432 when using the whole band data for modelling, leading to that not only the calculation becomes very slow but also the classification accuracy would be affected. Therefore, in this work, PCA was first used to reduce the dimensions of the whole spectrum data. According to the cross validation results, the first 10 PCs were selected as the input node m of the BP neural network, the output node n was set to 6 (i.e., 6 classes of gelatin gels), and the hidden layer node l was determined by the empirical formula of $l = \sqrt{m + n} + a$, where a is a constant between 0 and 10. During the training process, the network learning rate was set to 0.1, the number of training iterations was 1000, and the target error was 0.0001. The model was optimized by adjusting the hidden layer nodes, and an optimal three-layer BPNN model of 10 (input node)-12 (hidden layer node)-6 (output node) was obtained.

As described by Vitale et al. [18], in the case of classification involving m classes, a dummy binary-coded m -dimensional matrix can be used to describe the class that one

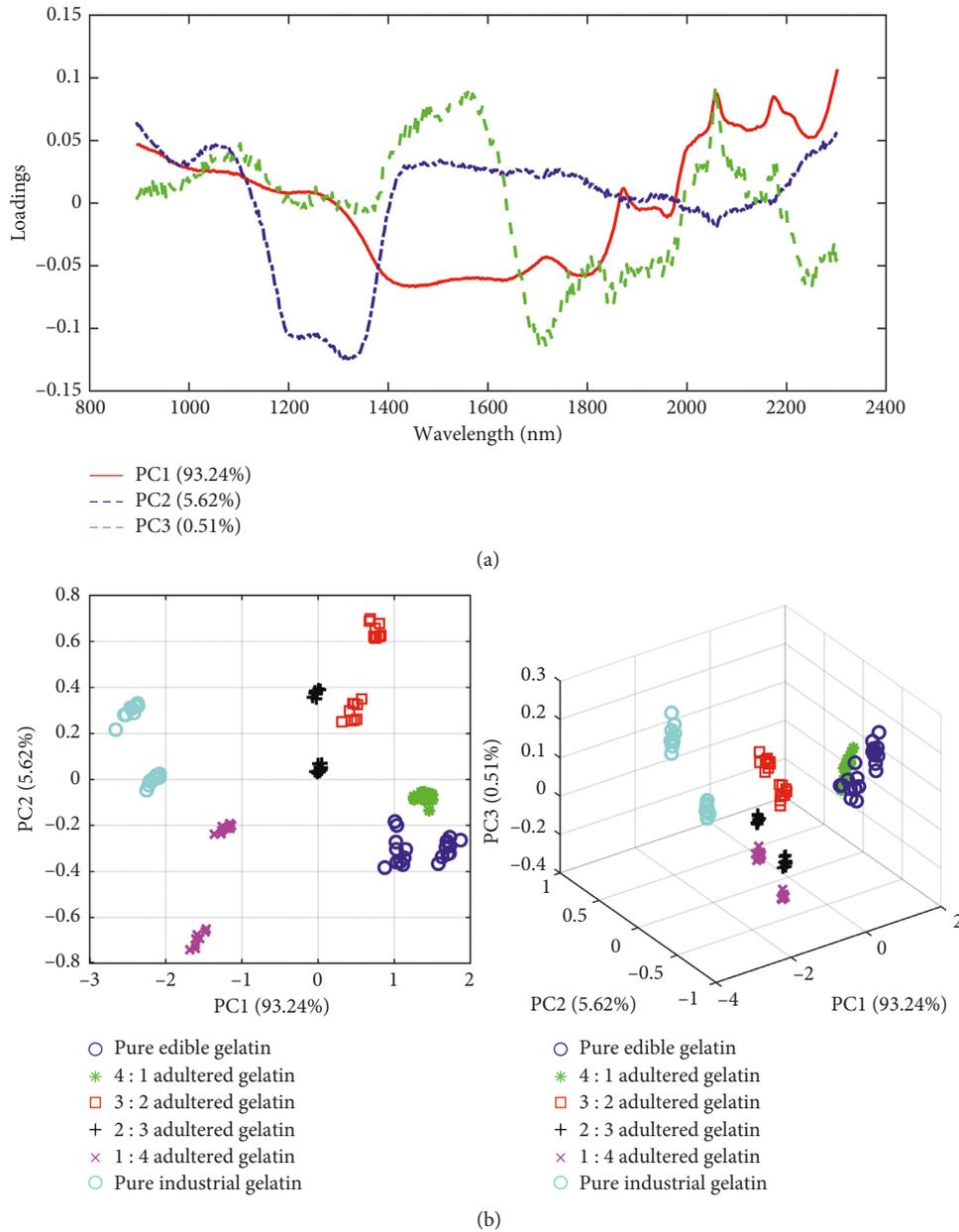


FIGURE 2: PCA loading plot: 2D and 3D score scatter plot of PC1, PC2, and PC3 for NIR spectra of gelatin samples.

TABLE 2: Predicting results of the LDA model in the training set and validation set.

Category	Training set		Validation set	
	Recognition rate	Rejection rate	Recognition rate	Rejection rate
Pure edible gelatin	100% (11/11)	100% (28/28)	90.91% (10/11)	100% (28/28)
4 : 1 adulterated gelatin	100% (7/7)	100% (32/32)	100% (7/7)	96.88% (31/32)
3 : 2 adulterated gelatin	100% (5/5)	100% (34/34)	100% (5/5)	100% (34/34)
2 : 3 adulterated gelatin	100% (7/7)	100% (32/32)	100% (7/7)	100% (32/32)
1 : 4 adulterated gelatin	100% (6/6)	100% (33/33)	100% (6/6)	100% (33/33)
Pure industrial gelatin	100% (3/3)	100% (36/36)	100% (3/3)	100% (36/36)

sample belongs to. In this study, samples belonging to pure edible gelatin (denoted by Class 1) were described by the vector [1 0 0 0 0], samples belonging to 4 : 1 adulterated gelatin (denoted by Class 2) by the vector [0 1 0 0 0],

samples belonging to 3 : 2 adulterated gelatin (denoted by Class 3) by the vector [0 0 1 0 0], samples belonging to 2 : 3 adulterated gelatin (denoted by Class 4) by the vector [0 0 0 1 0], samples belonging to 1 : 4 adulterated gelatin (denoted

TABLE 3: Predicting results of SIMCA model in the training set and validation set when the significance level is 0.05 and 0.1, respectively.

Category	PCs	$\alpha = 0.05$				$\alpha = 0.1$			
		Training set		Validation set		Training set		Validation set	
		Recognition rate (%)	Rejection rate (%)						
Pure edible gelatin	2	100	91	100	82	100	98	100	100
4:1 adulterated gelatin	7	100	100	100	100	100	100	100	100
3:2 adulterated gelatin	2	100	100	100	100	100	100	100	100
2:3 adulterated gelatin	4	100	100	100	100	100	100	100	100
1:4 adulterated gelatin	2	100	100	100	100	100	100	100	100
Pure industrial gelatin	4	100	100	100	100	100	100	100	100

by Class 5) by the vector [0 0 0 0 1 0], and samples belonging to pure industrial gelatin (denoted by Class 6) by the vector [0 0 0 0 0 1].

Under these assumptions, when the BPNN model was used for predicting 6 varieties of gelatin gels in the training set, all 75 samples were correctly identified without clustering overlap, producing a total recognition rate of 100%. As shown in Table 4, the predicting results of the BPNN model in the validation set show that only one sample (No. 23) was misclassified, which should be belonging to Class 4 and now incorrectly classified as Class 3, resulting in a recognition rate of 85.71% (6/7) for Class 4 and a rejection rate of 97.06% (33/34) for Class 3. Except for Class 4, the recognition rate of 100% for other classes was obtained, resulting in a recognition rate of 97.44% (38/39) for the validation set.

3.2.4. SVM Model. As we know, SVM is initially designed for binary classification. However, in recent years, SVM has been widely used for multiclass classification. To construct a SVM multiclass classifier, there are two common methods: one-versus-one (OVO) and one-versus-rest (OVR). In this work, we utilized the LIBSVM tool developed by Lin et al. to perform the SVM analysis, where LIBSVM tool adopted OVO method, thus 15 SVM classifiers were required for 6 groups of gelatin samples (i.e., k classes need $k(k-1)/2$ SVM classifiers). Radial basis function (RBF) was chosen as the kernel function for SVM classification. To obtain an optimal SVM model, the penalty parameter C and kernel parameter γ for the RBF function were optimized by using grid search (GS), genetic algorithm (GA), and particle swarm optimization (PSO) methods, the best C and γ were determined by the accuracy of cross validation (CV). The best optimization parameters and the predicting results of the SVM models are shown in Table 5, where the classified results of validation set (39 spectra) are summarized in a confusion matrix presented in Table 6, where each row represents actual class of samples and each column represents the classification by SVM models. The results show that the CV accuracy of these three optimization methods was 100%, and both the recognition rate and rejection rate in the training set and validation set

were 100%, indicating that GS, GA, and PSO methods have the same optimization ability to the SVM discrimination model. Moreover, the modelling results also demonstrate that the SVM model has a preferable discrimination ability to identify adulterated edible gelatin.

3.2.5. Performance Comparison of the Four Discrimination Models. Table 7 summarizes the predicting results of adulterated gelatin gels in the training set and validation set using LDA, SIMCA, BPNN, and SVM discrimination models, respectively. The results show that these four pattern recognition methods can well identify the adulterated gelatin gels, where the recognition rate in the training set was as high as 100%, and the recognition rate in the validation set was higher than 97%. For SIMCA and SVM models, the recognition rate both in the training set and validation set are 100%, which demonstrates that these two methods could completely recognize the 114 gelatin samples with 6 different adulteration levels. However, as displayed in Table 3, for the SIMCA model, when the significance level $\alpha = 0.05$, although the recognition rate for each class was 100%, five samples were overlapping which appeared in two classes. Only when the significance level $\alpha = 0.1$, all kinds of samples can be completely recognized. For the SVM model, the recognition rate and rejection rate both in the training set and verification set for each class were 100%, no matter which parameter optimization method was used. Therefore, among the four supervised pattern recognition methods, SVM is superior to the other three methods, which can fully identify the six classes of gelatin gels with different adulteration levels without clustering overlap.

4. Conclusion

Aiming at the problem of edible gelatin adulterated with industrial gelatin in the food industry, NIR absorption spectra of 114 gelatin gels from 6 kinds of adulteration gelatin were obtained with a compact NIR measuring setup; MSC, SG smoothing, and min-max normalization methods were first used for spectral preprocessing, and then PCA was utilized for the exploratory investigation, while PCA could

TABLE 4: Predicting results of BPNN model in the validation set.

No.	Real class	Predicted class	Result	No.	Actual species	Predicted class	Result	No.	Real class	Predicted class	Result
1	0 0 0 1 0 0	4	Correct	14	0 0 0 0 0 1	6	Correct	27	1 0 0 0 0 0	1	Correct
2	0 1 0 0 0 0	2	Correct	15	0 0 1 0 0 0	3	Correct	28	0 0 0 0 1 0	5	Correct
3	0 0 1 0 0 0	3	Correct	16	1 0 0 0 0 0	1	Correct	29	0 0 0 1 0 0	4	Correct
4	0 1 0 0 0 0	2	Correct	17	0 0 0 1 0 0	4	Correct	30	0 1 0 0 0 0	2	Correct
5	0 0 0 0 0 1	6	Correct	18	1 0 0 0 0 0	1	Correct	31	0 0 0 0 1 0	5	Correct
6	0 0 0 1 0 0	4	Correct	19	0 1 0 0 0 0	2	Correct	32	0 1 0 0 0 0	2	Correct
7	0 0 0 0 1 0	5	Correct	20	0 0 0 0 1 0	5	Correct	33	1 0 0 0 0 0	1	Correct
8	0 0 1 0 0 0	3	Correct	21	0 0 0 1 0 0	4	Correct	34	1 0 0 0 0 0	1	Correct
9	0 0 0 0 1 0	5	Correct	22	0 1 0 0 0 0	2	Correct	35	0 0 0 1 0 0	4	Correct
10	0 1 0 0 0 0	2	Correct	23	0 0 0 1 0 0	3	Error	36	1 0 0 0 0 0	1	Correct
11	0 0 1 0 0 0	3	Correct	24	1 0 0 0 0 0	1	Correct	37	0 0 1 0 0 0	3	Correct
12	0 0 0 0 1 0	5	Correct	25	0 0 0 0 0 1	6	Correct	38	1 0 0 0 0 0	1	Correct
13	1 0 0 0 0 0	1	Correct	26	1 0 0 0 0 0	1	Correct	39	1 0 0 0 0 0	1	Correct

Note. [1 0 0 0 0 0] represents class 1, i.e., pure edible gelatin ; [0 1 0 0 0 0] represents class 2, i.e., 4:1 adulterated gelatin ; [0 0 1 0 0 0] represents class 3, i.e., 3:2 adulterated gelatin ; [0 0 0 1 0 0] represents class 4, i.e., 2:3 adulterated gelatin ; [0 0 0 0 1 0] represents class 5, i.e., 1:4 adulterated gelatin ; [0 0 0 0 0 1] represents class 6, i.e., pure industrial gelatin.

TABLE 5: Optimization parameters and predicting results of SVM models based on three parameter optimization algorithms.

Parameter optimized methods	Best C	Best γ	CV accuracy	Recognition rate	
				Training set	Validation set
GS-SVM	0.3299	5.2780	100%	100% (75/75)	100% (39/39)
GA-SVM	1.1135	11.4377	100%	100% (75/75)	100% (39/39)
PSO-SVM	0.5283	15.7922	100%	100% (75/75)	100% (39/39)

TABLE 6: The classification results of SVM models in the validation set.

Number	Class 1	Class 2	Class 3	Class 4	Class 5	Class 6
Class 1	11	0	0	0	0	0
Class 2	0	7	0	0	0	0
Class 3	0	0	5	0	0	0
Class 4	0	0	0	7	0	0
Class 5	0	0	0	0	6	0
Class 6	0	0	0	0	0	3

TABLE 7: Performance comparison of 4 kinds of discrimination models.

Discrimination model	Recognition rate	
	Training set	Validation set
LDA	100% (75/75)	97.44% (38/39)
SIMCA	100% (75/75)	100% (39/39)
BPNN	100% (75/75)	97.44% (38/39)
SVM	100% (75/75)	100% (39/39)

not realize the obvious distinction of adulterated gelatin. Subsequently, the feasibility of adulterated gelatin identification was studied by building the LDA, SIMCA, BPNN, and SVM models, respectively. The modelling results show that all these four discrimination models present good recognition ability for the adulterated gelatin, the recognition rate

of which in the validation set is higher than 97%. Especially for the SIMCA and SVM models, the correct recognition rate reaches up to 100%. However, due to the possible clustering overlap between different classes when using the SIMCA model, the SVM model shows the preferable ability for the recognition of edible gelatin adulteration. Therefore, it is feasible and effective to use NIR spectroscopy combined with supervised pattern recognition methods for identifying edible gelatin adulteration rapidly and accurately.

Data Availability

The data used to support the findings of this study are available from the corresponding author upon request.

Conflicts of Interest

The authors declare that there are no conflicts of interest.

Acknowledgments

This work was financially supported by the China Post-doctoral Science Foundation (grant no. 2017M612399), the Science and Technology Project of Henan Province (grant nos. 182102110427 and 182102110250), the Science and Technology Innovation Project of Henan Agricultural University (grant no. KJCX2018A09), the National Natural Science Foundation of China (grant no. 31671581), and the Natural Science Foundation of Henan Province (grant no. 162300410143).

Supplementary Materials

In the supplementary files, Table S1 gives the list of used gelatin samples, Table S2 describes a total of 114 gelatin gelatin samples in the training set and validation set, Figure S1 displays the photographs of gelatin samples prepared for NIR spectral measurements, Table S3 gives the prediction results of LDA model in the validation set, the prediction results of the SIMCA model in the training set and validation set when the significance level $\alpha = 0.05$ are separately

displayed in Tables S4 and S5, and the prediction results of the SIMCA model in the training set and validation set when the significance level $\alpha = 0.1$ are shown in Tables S6 and S7, respectively. (*Supplementary Materials*)

References

- [1] H. Shah and F. Yusof, "Gelatin as an ingredient in food and pharmaceutical products: an Islamic perspective," *Advances in Environmental Biology*, vol. 8, pp. 774–780, 2014.
- [2] A. Jan, M. Azam, K. Siddiqui, A. Ali, I. Choi, and Q. Haq, "Heavy metals and human health: mechanistic insight into toxicity and counter defense system of antioxidants," *International Journal of Molecular Sciences*, vol. 16, no. 12, pp. 29592–29630, 2015.
- [3] X. Wang and Y. Wu, "Industrial gelatin poisons food; police crack ring," 2016, http://www.chinadaily.com.cn/china/2016-07/30/content_26278365.htm.
- [4] T. N. Azira, I. Amin, and Y. B. C. Man, "Differentiation of bovine and porcine gelatins in processed products via Sodium Dodecyl Sulphate-Polyacrylamide Gel Electrophoresis (SDS-PAGE) and principal component analysis (PCA) techniques," *International Food Research Journal*, vol. 19, pp. 1175–1180, 2012.
- [5] T. N. Azira, Y. B. C. Man, R. N. R. M. Hafidz, M. A. Aina, and I. Amin, "Use of principal component analysis for differentiation of gelatine sources based on polypeptide molecular weights," *Food Chemistry*, vol. 151, pp. 286–292, 2014.
- [6] H. Doi, E. Watanabe, H. Shibata, and S. Tanabe, "A reliable enzyme linked immunosorbent assay for the determination of bovine and porcine gelatin in processed foods," *Journal of Agricultural and Food Chemistry*, vol. 57, no. 5, pp. 1721–1726, 2009.
- [7] N. A. Tukiran, A. Ismail, S. Mustafa, and M. Hamid, "Determination of porcine gelatin in edible bird's nest by competitive indirect ELISA based on anti-peptide polyclonal antibody," *Food Control*, vol. 59, pp. 561–566, 2016.
- [8] N. A. Tukiran, A. Ismail, S. Mustafa, and M. Hamid, "Development of antipeptide enzyme-linked immunosorbent assay for determination of gelatin in confectionery products," *International Journal of Food Science & Technology*, vol. 51, no. 1, pp. 54–60, 2015.
- [9] G. Zhang, T. Liu, Q. Wang et al., "Mass spectrometric detection of marker peptides in tryptic digests of gelatin: a new method to differentiate between bovine and porcine gelatin," *Food Hydrocolloids*, vol. 23, no. 7, pp. 2001–2007, 2009.
- [10] M. T. Yilmaz, Z. Kesmen, B. Baykal et al., "A novel method to differentiate bovine and porcine gelatins in food products: NanoUPLC-ESI-Q-TOF-MSE based data independent acquisition technique to detect marker peptides in gelatin," *Food Chemistry*, vol. 141, no. 3, pp. 2450–2458, 2013.
- [11] M. I. Azilawati, D. M. Hashim, B. Jamilah, and I. Amin, "RP-HPLC method using 6-aminoquinolyl-N-hydroxysuccinimidyl carbamate incorporated with normalization technique in principal component analysis to differentiate the bovine, porcine and fish gelatins," *Food Chemistry*, vol. 172, pp. 368–376, 2015.
- [12] H. Cai, X. Gu, M. S. Scanlan, D. H. Ramatlapeng, and C. R. Lively, "Real-time PCR assays for detection and quantitation of porcine and bovine DNA in gelatin mixtures and gelatin capsules," *Journal of Food Composition and Analysis*, vol. 25, no. 1, pp. 83–87, 2012.
- [13] S. A. Mutalib, N. M. Muin, A. Abdullah et al., "Sensitivity of polymerase chain reaction (PCR)-southern hybridization and conventional PCR analysis for Halal authentication of gelatin capsules," *LWT-Food Science and Technology*, vol. 63, no. 1, pp. 714–719, 2015.
- [14] H. I. A. Amqizal, H. A. Al-Kahtani, E. A. Ismail, K. Hayat, and I. Jaswir, "Identification and verification of porcine DNA in commercial gelatin and gelatin containing processed foods," *Food Control*, vol. 78, pp. 297–303, 2017.
- [15] B. Jannat, K. Ghorbani, H. Shafieyan et al., "Gelatin specification using real-time PCR and analysis of mass spectrometry-based proteomics datasets," *Food Control*, vol. 87, pp. 79–87, 2018.
- [16] D. M. Hashim, Y. B. C. Man, R. Norakasha, M. Shuhaimi, Y. Salmah, and Z. A. Syahariza, "Potential use of Fourier transform infrared spectroscopy for differentiation of bovine and porcine gelatins," *Food Chemistry*, vol. 118, no. 3, pp. 856–860, 2010.
- [17] N. Cebi, M. Z. Durak, O. S. Toker, O. Sagdic, and M. Arici, "An evaluation of Fourier transforms infrared spectroscopy method for the classification and discrimination of bovine, porcine and fish gelatins," *Food Chemistry*, vol. 190, pp. 1109–1115, 2016.
- [18] R. Vitale, M. Bevilacqua, R. Bucci, A. D. Magrì, A. L. Magrì, and F. Marini, "A rapid and non-invasive method for authenticating the origin of pistachio samples by NIR spectroscopy and chemometrics," *Chemometrics and Intelligent Laboratory Systems*, vol. 121, pp. 90–99, 2013.
- [19] X. Ding, Y. Ni, and S. Kokot, "NIR spectroscopy and chemometrics for the discrimination of pure, powdered, purple sweet potatoes and their samples adulterated with the white sweet potato flour," *Chemometrics and Intelligent Laboratory Systems*, vol. 144, pp. 17–23, 2015.
- [20] F. Mabood, F. Jabeen, M. Ahmed et al., "Development of new NIR-spectroscopy method combined with multivariate analysis for detection of adulteration in camel milk with goat milk," *Food Chemistry*, vol. 221, pp. 746–750, 2017.
- [21] L.-Q. Hu, C.-L. Yin, and Z.-P. Zeng, "Detection of adulteration in acetonitrile using near infrared spectroscopy coupled with pattern recognition techniques," *Spectrochimica Acta Part A: Molecular and Biomolecular Spectroscopy*, vol. 151, pp. 34–39, 2015.
- [22] P. Liu, J. Wang, Q. Li, J. Gao, X. Tan, and X. Bian, "Rapid identification and quantification of Panax notoginseng with its adulterants by near infrared spectroscopy combined with chemometrics," *Spectrochimica Acta Part A: Molecular and Biomolecular Spectroscopy*, vol. 206, pp. 23–30, 2019.
- [23] H. Zhang, S. Wang, K. Chang et al., "Optical characterization of paper aging based on laser-induced fluorescence (LIF) spectroscopy," *Applied Spectroscopy*, vol. 72, no. 6, pp. 913–920, 2018.
- [24] G. McLachlan, *Discriminant Analysis and Statistical Pattern Recognition*, John Wiley & Sons Inc., New York, NY, USA, 2005.
- [25] S. Wold and M. Sjöström, "SIMCA: a method for analysing chemical data in terms of similarity and analogy," in *Chemometrics, Theory and Application*, American Chemical Society, Washington, DC, USA, pp. 243–282, 52nd edition, 1977.
- [26] B. Mertens, M. Thompson, and T. Fearn, "Principal component outlier detection and SIMCA: a synthesis," *The Analyst*, vol. 119, pp. 2777–2784, 1994.
- [27] C. M. Bishop, *Neural Networks for Pattern Recognition*, Clarendon Press, Oxford, UK, 1995.
- [28] C. Cortes and V. Vapnik, "Support-vector Networks," *Machine Learning*, vol. 20, no. 3, pp. 273–297, 1995.

- [29] C.-C. Chang and C.-J. Lin, "Libsvm," *ACM Transactions on Intelligent Systems and Technology*, vol. 2, no. 3, pp. 1–27, 2011.
- [30] L. Cséfalvayová, M. Pelikan, I. Kralj Cigić, J. Kolar, and M. Strlič, "Use of genetic algorithms with multivariate regression for determination of gelatine in historic papers based on FT-IR and NIR spectral data," *Talanta*, vol. 82, no. 5, pp. 1784–1790, 2010.
- [31] A. Duconseille, D. Andueza, F. Picard, V. Santé-Lhoutellier, and T. Astruc, "Molecular changes in gelatin aging observed by NIR and fluorescence spectroscopy," *Food Hydrocolloids*, vol. 61, pp. 496–503, 2016.
- [32] A. Duconseille, D. Andueza, F. Picard, V. Santé-Lhoutellier, and T. Astruc, "Variability in pig skin gelatin properties related to production site: a near infrared and fluorescence spectroscopy study," *Food hydrocolloids*, vol. 63, pp. 108–119, 2017.



Hindawi

Submit your manuscripts at
www.hindawi.com

