

Application of Near-Infrared Spectroscopy and Fuzzy Improved Null Linear Discriminant Analysis for Rapid Discrimination of Milk Brands

Xiaohong Wu ^{1,2} , Yiheng Fang ¹, Bin Wu ^{3,*} and Man Liu ¹

¹ School of Electrical and Information Engineering, Jiangsu University, Zhenjiang 212013, China; wxh419@ujs.edu.cn (X.W.); 17857578515@163.com (Y.F.); 18186214786@163.com (M.L.)

² High-Tech Key Laboratory of Agricultural Equipment and Intelligence of Jiangsu Province, Jiangsu University, Zhenjiang 212013, China

³ Department of Information Engineering, Chuzhou Polytechnic, Chuzhou 239000, China

* Correspondence: wubin@chzc.edu.cn

Abstract: The quality of milk is tightly linked to its brand. A famous brand of milk always has good quality. Therefore, this study seeks to design a new fuzzy feature extraction method, called fuzzy improved null linear discriminant analysis (FiNLDA), to cluster the spectra of collected milk for identifying milk brands. To elevate the classification accuracy, FiNLDA was applied to process the near-infrared (NIR) spectra of milk acquired by the portable near-infrared spectrometer. The principal component analysis and Savitzky–Golay (SG) filtering algorithm were employed to lower dimensionality and eliminate noise in this system, respectively. Thereafter, improved null linear discriminant analysis (iNLDA) and FiNLDA were applied to attain the discriminant information of the NIR spectra. At last, the K-nearest neighbor classifier was utilized for assessing the performance of the identification system. The results indicated that the maximum classification accuracies of LDA, iNLDA and FiNLDA were 74.7%, 88% and 94.67%, respectively. Accordingly, the portable NIR spectrometer in combination with FiNLDA can classify milk brands correctly and effectively.

Keywords: milk; near-infrared spectroscopy; improved null linear discriminant analysis; Savitzky–Golay filtering; K-nearest neighbor



Citation: Wu, X.; Fang, Y.; Wu, B.; Liu, M. Application of Near-Infrared Spectroscopy and Fuzzy Improved Null Linear Discriminant Analysis for Rapid Discrimination of Milk Brands. *Foods* **2023**, *12*, 3929.

<https://doi.org/10.3390/foods12213929>

Academic Editors: Barbara Giussani and Mircea Oroian

Received: 25 September 2023

Revised: 18 October 2023

Accepted: 25 October 2023

Published: 26 October 2023



Copyright: © 2023 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

As an essential nutritional source for numerous people [1], milk encompasses multiple nutritional or non-nutritional elements, such as functional and bioactive substances, lipids [2], and proteins [3], as well as minerals [4], vital amino acids, and lactose [5]. Although the milk of all mammals possesses the same primary components, namely carbohydrates, water, minerals, proteins, vitamins, and fats, the milk of cud-chewing and non-cud-chewing animals varies substantially in the contents of these components [6]. The composition of milk is also affected by the frequency and stage of lactation, heredity, somatic cell count, animal diet, the treatment of milk, and seasonal variations [7–11].

However, milk is a major object of adulteration because of its enormous economic benefits. The nutritional assessment and adulteration detection of milk has become increasingly challenging due to the features of milk, such as special composition, the broad range of animal origins, and many influencing factors for nutrition. The evolution of food omics in recent years has enabled the increasing application of various techniques for rapidly screening or selectively confirming milk quality and authenticity.

Near-infrared (NIR) spectroscopy coupled with multivariate processing techniques has been extensively analyzed in recent years. This technology has been significantly applied in several industries, like the agricultural product industry [12,13], the pharmaceutical factory [14] and the petrochemical industry [15,16]. NIR has become a pivotal assessment

method for milk quality [17,18] because it is an undamaged, green and rapid method. More importantly, NIR analysis is free of chemical reagents in line with the principle of green chemistry. In addition to these advantages, NIR spectroscopy has been validated to be utilized for the good analysis of food and agricultural products because NIR mainly records the frequency-doubled and frequency-combined absorption of the hydrogen-containing group X-H (X = C, N, O) vibration [19]. NIR can be used in identifying food varieties, and furthermore, the analytical methods require discriminant analysis and classifiers. The portable NIR instrument has multiple advantages, such as the ability of emergency analysis, a short test time, an undamaged sample, a small sample size, and easy maintenance, but the spectrum detection range is narrow [20].

Given the serious overlap of spectra in the NIR spectrum region, if NIR spectra are directly classified, the classification accuracy will be low. To improve the classification accuracy, NIR spectra should be processed by a feature extraction method before classification. Improved null linear discriminant analysis (iNLDA) is an improved feature extraction method based on null linear discriminant analysis (NLDA) [21]. iNLDA and NLDA have similar capability in most cases, but iNLDA further cuts down the computational cost of NLDA. Nevertheless, for the within-class scatter matrix constructed by the low-dimensional data, its null space is empty, and neither NLDA nor iNLDA is applicable [22]. In order to obtain a faster, more accurate and more flexible algorithm to process the NIR spectra of collected milk, this paper presented a novel fuzzy feature extraction method, called improved null linear discriminant analysis (FiNLDA). And a combination of portable NIR spectrometer and FiNLDA was designed to identify milk brands.

2. Materials and Methods

2.1. Sample Preparation

In this study, five brands (Guangming, Mengniu, Telunsu, Yili, Jindian) of milk samples came from the local supermarket in China. For the same brand of milk samples, they came from the same manufacturer, and their production batch and production date are the same; for different brands of milk samples, their production dates are close. There were 60 samples for each brand, totaling 300 samples. Thereafter, all milk samples were categorized into training and test samples according to a certain proportion. The milk samples met the following requirements: milk capacity (250 mL), packaging (carton/plastic shell) and in the shell life.

2.2. NIR Spectra Collection

The NIR spectral data of milk samples were acquired using the NIR-M-R2 spectrometer made by Shenzhen Pynect limited corporation, Shenzhen, China. The wavelength range of the spectrometer ranges from 900 nm to 1700 nm, and its resolution is 10 nm. The NIR data were collected at about 25 °C with relative humidity of 50–60% throughout the collection process. The spectrometer should be warmed up for one hour before collecting near-infrared spectral data. The collected near-infrared spectrum of milk is the 228-dimensional data. Each milk sample was scanned three times, and the final datum was the average of the three test results. The NIR spectra of milk samples are shown in Figure 1.

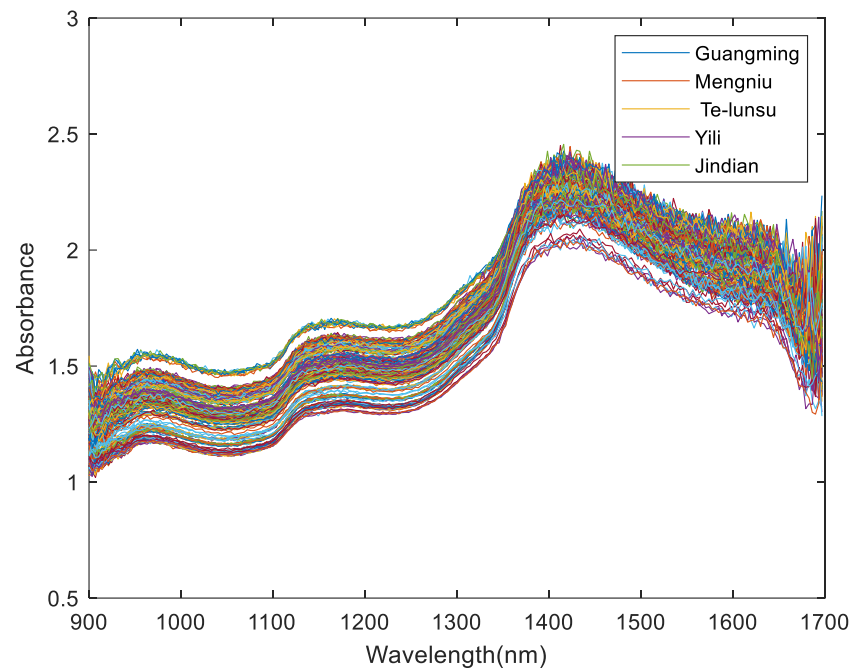


Figure 1. The raw spectra of milk.

2.3. NIR Spectra Preprocessing

The original spectrum is greatly affected by physical conditions. The data shown in Figure 1 were mixed with a noise signal in addition to the sample characteristics [23]. In order to filter out the noise signal, the NIR spectra needed to be preprocessed, and the Savitzky–Golay (SG) filtering method was applied to preprocess the NIR spectra [24]. Regarding the SG filter, the `sgolayfilt(X, order, frame)` function in MATLAB was performed to preprocess the data. This experiment defined order as 1 and frame as 11. The functions of setting these parameters are to remove scattering, cut down the influence of diffuse reflection, reduce random errors, remove the redundant data, etc. [22,25]. After preprocessing, the NIR spectra are shown in Figure 2.

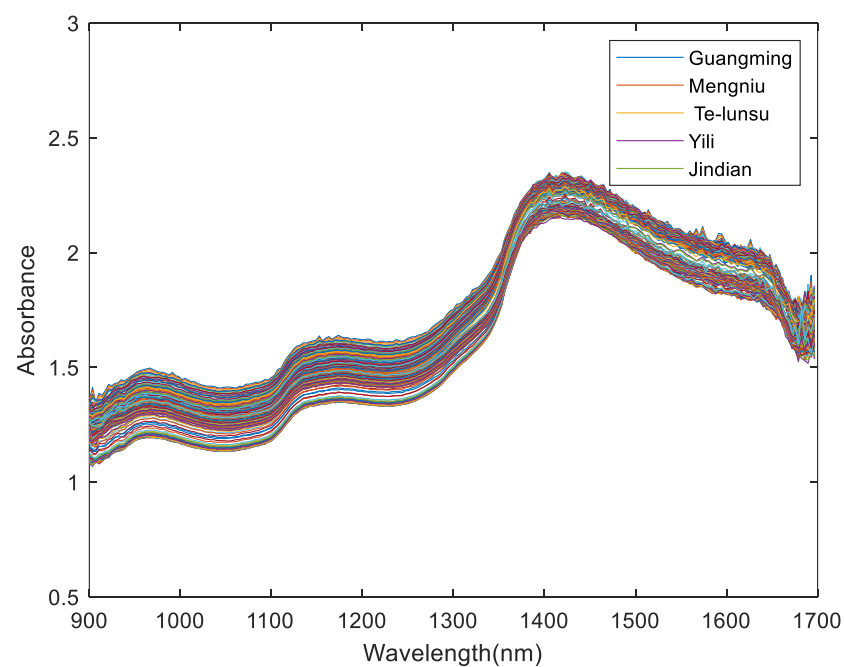


Figure 2. The preprocessed spectra of milk.

2.4. Improved Null Linear Discriminant Analysis

The procedure of iNLDA is described as follows [22]:

- (1) Build matrices H_t, H_b, H_w by the training data (containing n data points in \mathbb{R}^m);
- (2) Performing singular value decomposition on the matrix $H_t, H_t = U_1 \Sigma_r V_1, \Sigma_r \in \mathbb{R}^{r \times r}, U_1 \in \mathbb{R}^{m \times r}, U_2 \in \mathbb{R}^{m \times (n-r)}, r = \text{rank}(S_t)$;
- (3) Construct the matrix $\widetilde{S}_w = U_1^T S_w U_1$;
- (4) Perform the eigendecomposition of matrix \widetilde{S}_w . The matrix W is constructed by the eigenvectors associated with the zero eigenvalues;
- (5) Define the matrix $G = U_1 W$. Here, G is the feature projection matrix of iNLDA.

2.5. Fuzzy Improved Null Linear Discriminant Analysis

The procedure of FiNLDA is described as follows (For the calculation of the initial fuzzy membership degree, see Formula (1) in the Ref. [26]):

- (1) Build matrices H_{ft}, H_{fb}, H_{fw} by the training data (containing n data points in \mathbb{R}^p):

$$H_{ft} = \left[\sqrt{\sum_{j=1}^c u_{1j}^m (x_1 - \bar{x})}, \sqrt{\sum_{j=1}^c u_{2j}^m (x_2 - \bar{x})}, \dots, \sqrt{\sum_{j=1}^c u_{nj}^m (x_n - \bar{x})} \right] \quad (1)$$

$$H_{fb} = \left[\sqrt{(U_1^f)^T e (v_1 - \bar{x})}, \sqrt{(U_2^f)^T e (v_2 - \bar{x})}, \dots, \sqrt{(U_c^f)^T e (v_c - \bar{x})} \right] \quad (2)$$

$$H_{fw} = [A_1, A_2, \dots, A_c] \quad (3)$$

Here, U_j^f and A_j in (2) and (3), are defined as

$$U_j^f = [u_{1j}^m, u_{2j}^m, \dots, u_{nj}^m]^T \quad (4)$$

$$A_j = \left[\sqrt{u_{1j}^m (x_1 - v_j)}, \sqrt{u_{2j}^m (x_2 - v_j)}, \dots, \sqrt{u_{nj}^m (x_n - v_j)} \right] \quad (5)$$

where \bar{x} is the mean of all samples $\bar{x} = \frac{1}{n} \sum_{i=1}^n x_i = \frac{1}{n} X e, e = [1, 1, \dots, 1]^T \in \mathbb{R}^n$.

- (2) Performing singular value decomposition on the matrix $H_{ft}, H_{ft} = U_1 \Sigma_r V_1, U_1 \in \mathbb{R}^{p \times r}, U_2 \in \mathbb{R}^{p \times (n-r)}, r = \text{rank}(S_{ft})$;
- (3) Construct the matrix $\widetilde{S}_{fw} = U_1^T S_{fw} U_1$;
- (4) Perform the eigendecomposition of matrix \widetilde{S}_{fw} . The matrix W is constructed by the eigenvectors associated with the zero eigenvalues;
- (5) Define the matrix $G = U_1 W$. Here, G is the feature projection matrix of FiNLDA.

2.6. K-Nearest Neighbor

One superiority of the K-nearest neighbor (KNN) method is its simple calculation, but it does not affect its classification effect. In some experiments, KNN can even achieve better classification results than other more complex classifiers. For KNN, the K-nearest samples closest to the unknown object are chosen and the majority rule is applied: the unknown object is categorized into the class to which most of the K samples belong to. The selection of K is performed by calculating the predictive power for different values of K, with small values of K (3 or 5) generally being preferred [27].

2.7. Software

In this experiment, MATLAB R2022a (The MathWorks, Co., Ltd., Portolla Valley, CA, USA) was utilized to deal with the data and calculate the classification accuracy.

3. Results

3.1. Dimensional Reduction by PCA

In this experiment, the NIR spectral data of our milk samples were adopted to construct a data matrix of 300×228 , where 300 is the sample number and 228 is the sample dimension. Since the data matrix contains feature information irrelevant to the analysis, this experiment used principal component analysis (PCA) to reduce the dimension and redundancy of the data. After PCA processing, the reduced dimension matrix including original features, orthogonal features and integrated features can be obtained.

The number of principal components (PCs) can markedly impact the classification accuracy of the recognition system. Accordingly, it is highly critical for selecting the proper number of PCs to improve the classification accuracy [28].

The number of principal components is usually selected based on the cumulative contribution rate. When the rate reaches 90%, it proves that the principal components contain most of the effective information. Since the first 20 principal components constitute 99% of the total variance, they retain the feature data of the near-infrared spectral data, and eliminate the redundant data. After 20 feature vectors were calculated by PCA, the 228-dimensional spectra were projected into these vectors to produce the 20-dimensional data. Furthermore, a three-dimensional feature space constructed by the first three principal components was established to observe the near-infrared spectral data of milk. The three-dimensional display of the training data after PCA processing is shown in Figure 3. After PCA, the 300 milk samples were divided into a training set (45 training samples per brand, 225 in total) and a test set (15 test samples per brand, 75 in total).

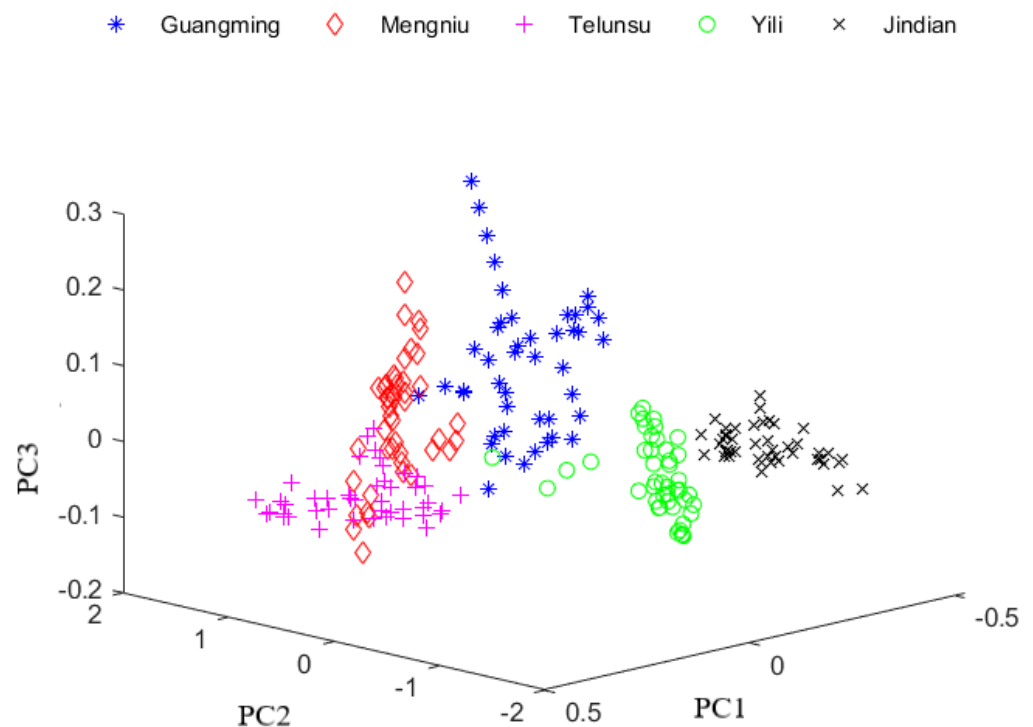


Figure 3. Data distribution of training set under S-G filter + PCA.

The classification of PCA + LDA, PCA + iNLDA and PCA + FiNLDA for milk brands were introduced in the following sections.

3.2. Discriminant Feature Extraction by LDA

After PCA reduced the dimension of the NIR data to 20 dimensions, linear discriminant analysis (LDA) was applied to extract the discriminant features from the training set. For LDA, the number of eigenvectors and eigenvalues is usually the category number minus one. Therefore, the eigenvalues of four features were computed and listed: $\lambda_1 = 113.2$, $\lambda_2 = 51.6$, $\lambda_3 = 21.9$, $\lambda_4 = 4.8$. The twenty-dimensional test data were projected onto the first three eigenvectors (LDV1, LDV2 and LDV3) of LDA to generate three-dimensional data in Figure 4. Figure 4 illustrated the scattergram of data points performed by PCA + LDA. This figure showed that the data points of Guangming and Mengniu have crossed parts, which are difficult to be distinguished. It was clear that PCA + LDA has a certain classification function for data, but it could not clearly depict the classification boundary. Therefore, this experiment used the improved null linear discriminant analysis (iNLDA) to improve the data distribution for classification.

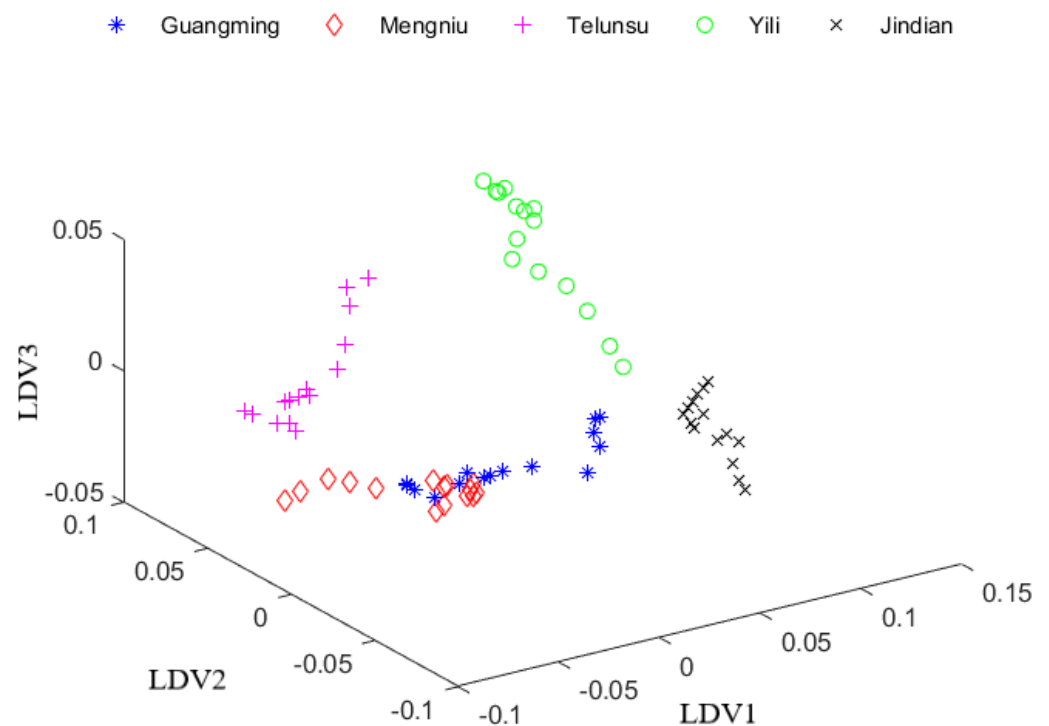


Figure 4. The data distribution of test set after PCA + LDA.

3.3. Discriminant Feature Extraction by iNLDA

iNLDA can increase the compactness of data points in the same class and separation of data points in different classes. The class of data points can be accurately determined by classifiers, like the KNN classifier. After iNLDA processed the training data to generate four discriminant vectors (iNLDV1, iNLDV2, iNLDV3 and iNLDV4), the 20-dimensional test samples were transformed into the 4-dimensional data, whose first three-dimensional data were shown in Figure 5. Five kinds of milk samples are distributed clearly. However, the data distributions of Guangming, Mengniu and Telunsu were a little close. Therefore, there might still be problems such as misclassification and identification failure. Nevertheless, compared with LDA, the recognition accuracy of iNLDA has been greatly improved.

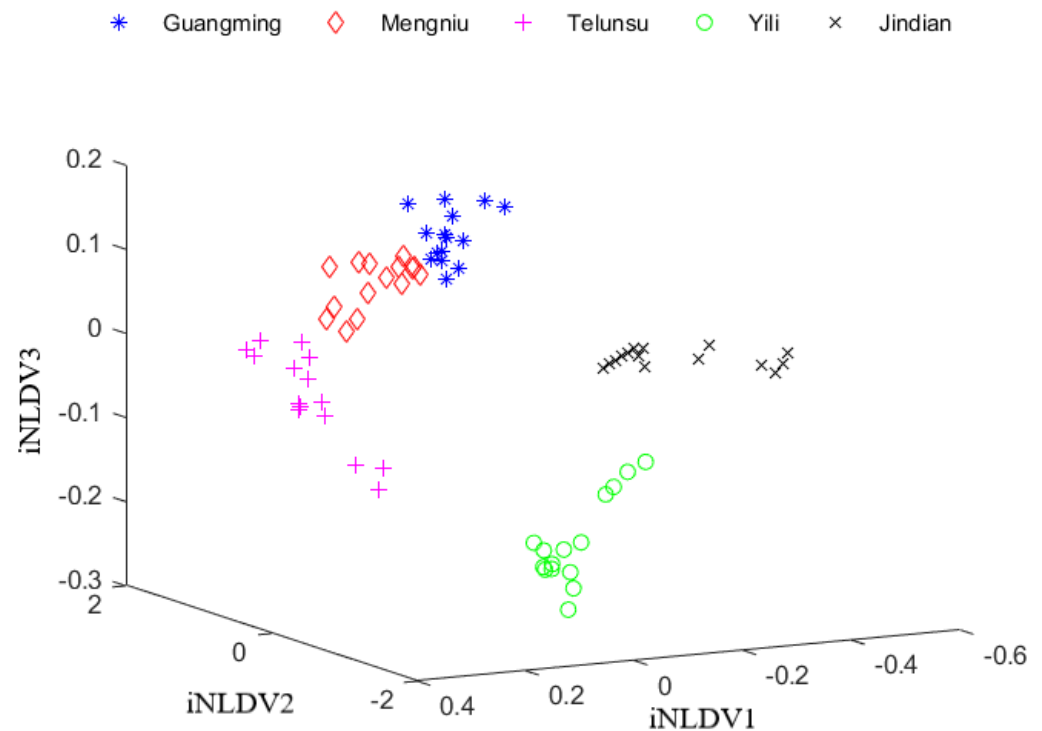


Figure 5. The data distribution of test set after PCA + iNLDA.

3.4. Discriminant Feature Extraction by FiNLDA

To reduce the limitation of iNLDA, FiNLDA was introduced to improve the recognition accuracy and avoid the misclassification. This experiment introduced fuzzy membership to initialize the training samples to improve the performance of processing overlapped data. The parameters of FiNLDA were set: the fuzzy index $m = 2.0$, the number of class $c = 5$. The fuzzy membership degrees were calculated by the equation described in fuzzy c-means (FCM) clustering. The initial cluster centers of FiLDA were the mean values of each brand of milk samples, and they are shown in Equation (6) with application of the initial cluster center model to classification and calculation accuracy.

The initial fuzzy membership values of FiNLDA are shown in Figure 6. The abscissa represents the sample data and the ordinate stands for the fuzzy membership value. This experiment involved five different brands of milk samples. Hence, there are five different little figures. Each little figure represents a brand of milk, i.e., Guangming, Mengniu, Telunsu, Yili and Jindian. When the fuzzy membership degree of the i -th sample u_{ij} is the largest in the j -th class, it can be determined that the i -th sample belongs to the j -th class.

After PCA + FiNLDA, the data distribution is shown in Figure 7. From the fuzzy membership and data distribution, it could be seen that different brands of milk samples can be separated well by the FiNLDA algorithm.

$$v^{(0)} = \begin{pmatrix} v_1^{(0)} \\ v_2^{(0)} \\ v_3^{(0)} \\ v_4^{(0)} \\ v_5^{(0)} \end{pmatrix} = \begin{pmatrix} 1.0527 & -0.2374 & -0.0451 & \cdots & 0.0016 \\ 0.8488 & 0.1277 & -0.0212 & \cdots & -0.0022 \\ 0.3955 & 0.3278 & 0.0478 & \cdots & 0.0020 \\ -1.7051 & 0.1550 & -0.0257 & \cdots & -0.0018 \\ -0.6238 & -0.3968 & 0.0360 & \cdots & 0.0017 \end{pmatrix}_{5 \times 20} \quad (6)$$

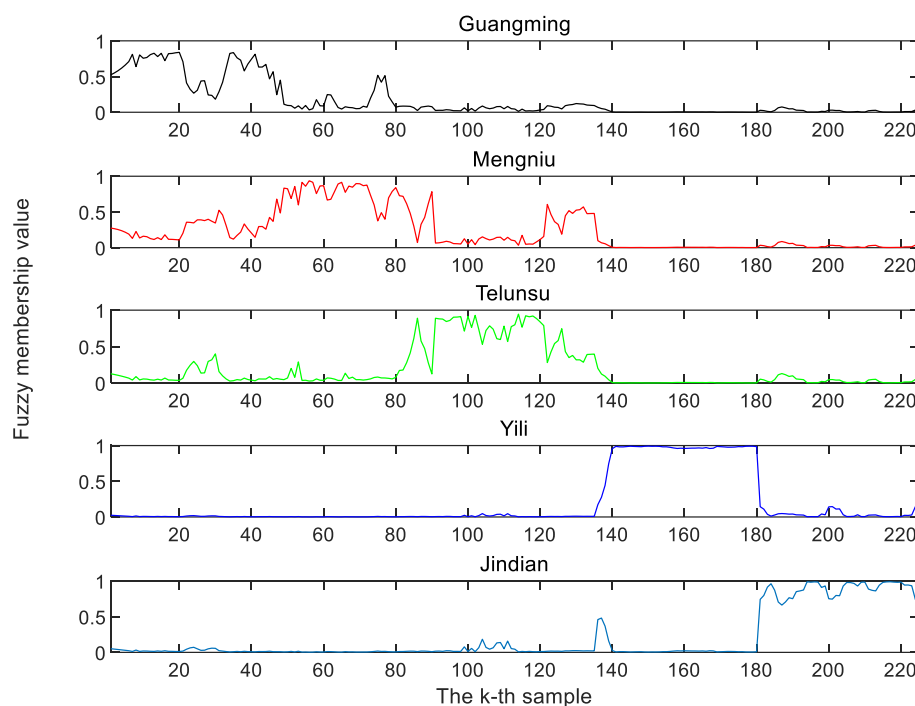


Figure 6. Fuzzy membership of FiNLDA.

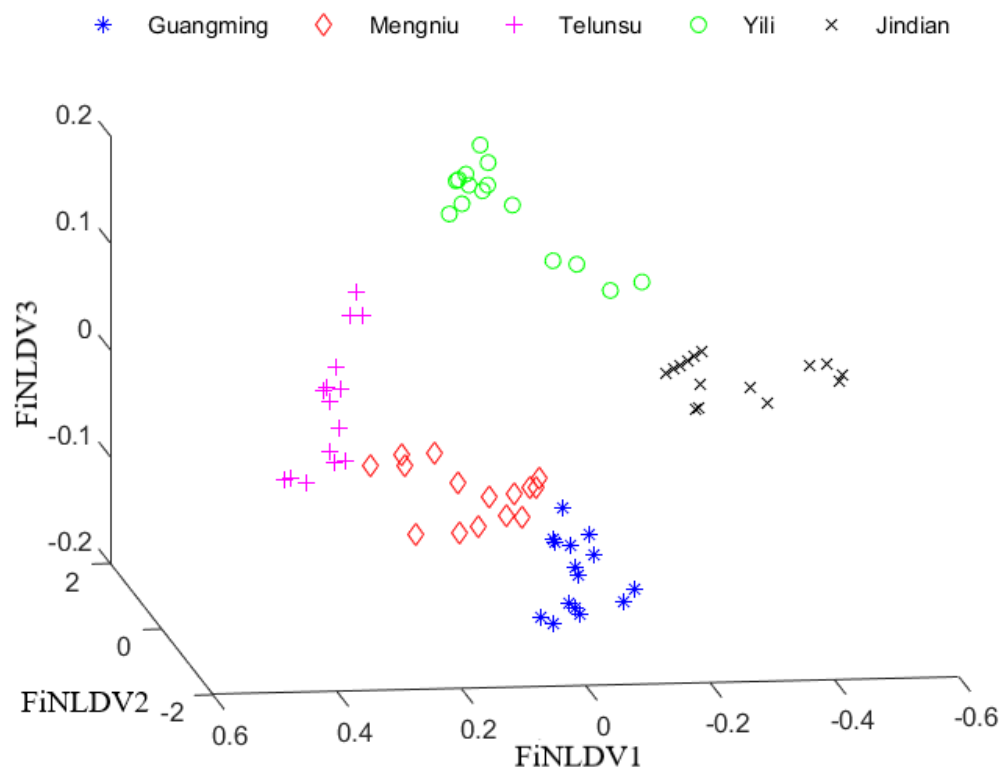


Figure 7. The data distribution after PCA + FiNLDA.

3.5. Classification Results

The K-nearest neighbor (KNN) classifier has many advantages. For example, it is simple and effective, and it is suitable for cross-domain samples and automatic classification of large samples. This experiment can use its characteristics to classify the milk samples. This study used KNN to compare the classification accuracies of different algorithms and different values of K. MATLAB was applied to calculate the accuracy of the KNN

classifier. Take $K = 3$ for example. The experimental results illustrated that the identification accuracies of LDA, iNLDA and FiNLDA are 74.7%, 88% and 93.3%, respectively. It turned out that FiNLDA has a good classification effect on milk brands.

3.6. Classify Accuracy under Different Values of K

It is well known that different values of K for a KNN classifier can affect the classification accuracy. If this experiment chooses a smaller K value, our model will become complex and easily overfitted. Additionally, the selection of a larger K value in our experiment represents prediction with the training data in a larger neighborhood, thus resulting in the wrong prediction since the training examples that are far away from the input examples (not similar) will also be involved in the prediction. An increase in the K value indicates that the model is simplified. Therefore, it is advisable to find the optimal K value. This study could use the different K values of the KNN classifier and compare the classification accuracy of the classifier. On this basis, this study calculated the classification accuracy under different K values to screen the optimal K value to achieve the classification goal. The classification results were shown in Figure 8. After the comparison, this experiment found when the K value reached to 3, the classification accuracy was the highest. Therefore, this experiment chose 3 as the K value of the KNN classifier.

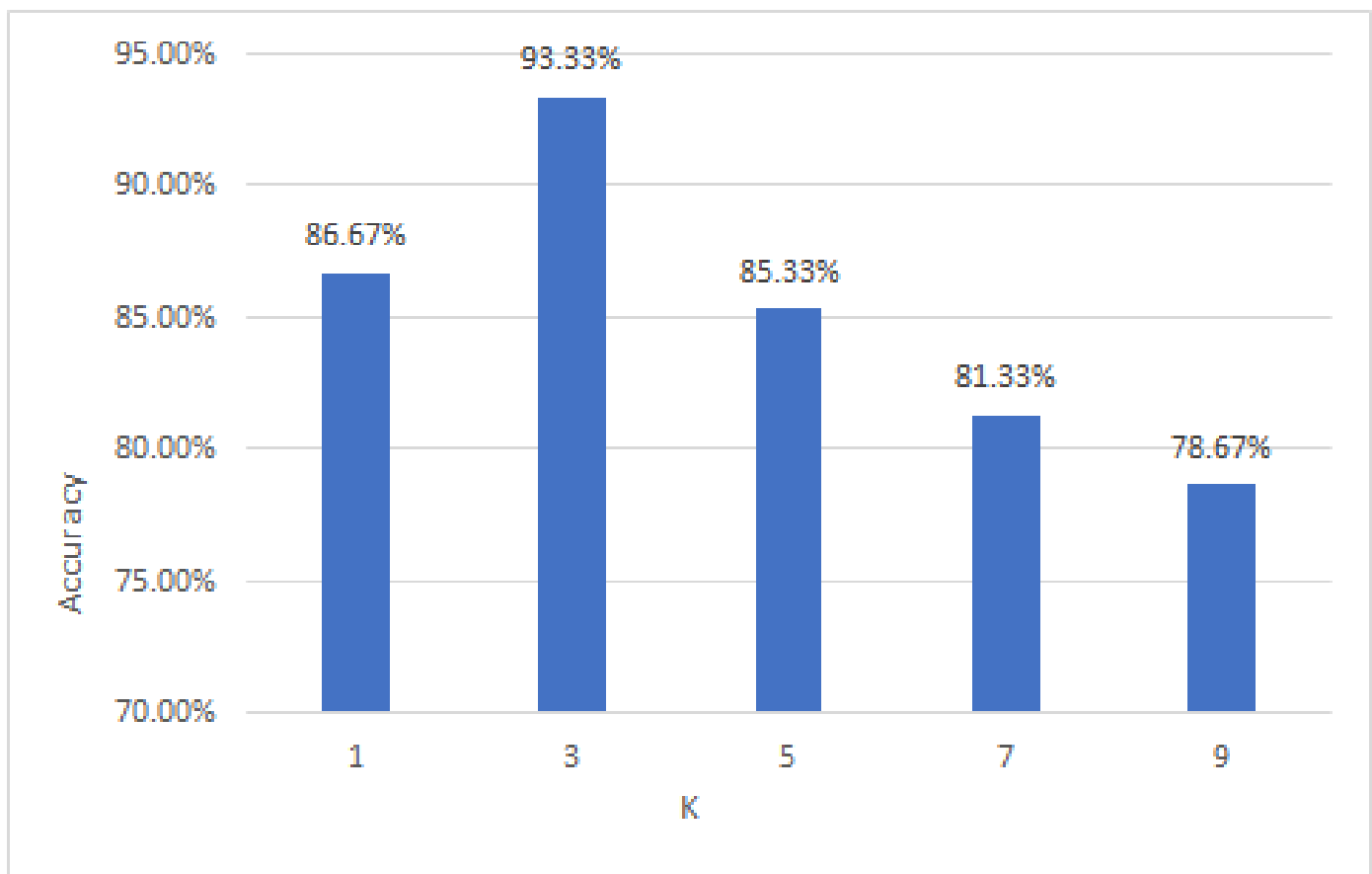


Figure 8. Classification accuracy under different K values of KNN.

4. Discussion

In order to identify the brand of milk samples correctly, this study aimed to extract discriminant features effectively from the NIR spectra of milk samples by proposing a novel fuzzy feature extraction method, i.e., FiNLDA, which was a combination of a fuzzy set and iNLDA. The NIR spectra of milk samples contained a noise signal and they were overlapped seriously, and this made it difficult to identify the spectra of different brands

of milk samples. As the “hard” feature extraction method, LDA and iNLDA are not satisfactory in the results of processing spectral data. FiNLDA, as a “soft” feature extraction method, demonstrated the excellent ability to process spectral data. The experiment showed that FiNLDA achieved the highest classification accuracy of 93.33%, which was higher than LDA and iNLDA. Furthermore, the FiNLDA-based classification model was tested with different K values of KNN, and the classification accuracies were higher than 85% when K was 1, 3 and 5, respectively.

As a fuzzy discriminant analysis, the FiNLDA algorithm used the fuzzy between/within scatter matrix to compute eigen decomposition, and the weight index m in the matrix has a powerful influence on the classification accuracy. If m becomes larger, the u_{ij}^m is smaller and “fuzzier”. When $m \rightarrow +\infty$, the fuzzy membership $u_{ij} \rightarrow 1/c$ [29], and usually $0 \leq u_{ij} < 1$, thus, $u_{ij}^m \rightarrow 0$. On the other hand, when $m \rightarrow 1$, the fuzzy between/within scatter matrix becomes the “hard” between/within scatter matrix. The choice of a suitable m is still an open problem and has no theoretical basis. This study tried different m weight exponents to figure out the classification accuracies, which were shown in Figure 9. From Figure 9, the classification accuracy of FiNLDA reached the highest of 94.67% when the optimal value of m was 1.5, 1.6 and 1.8.

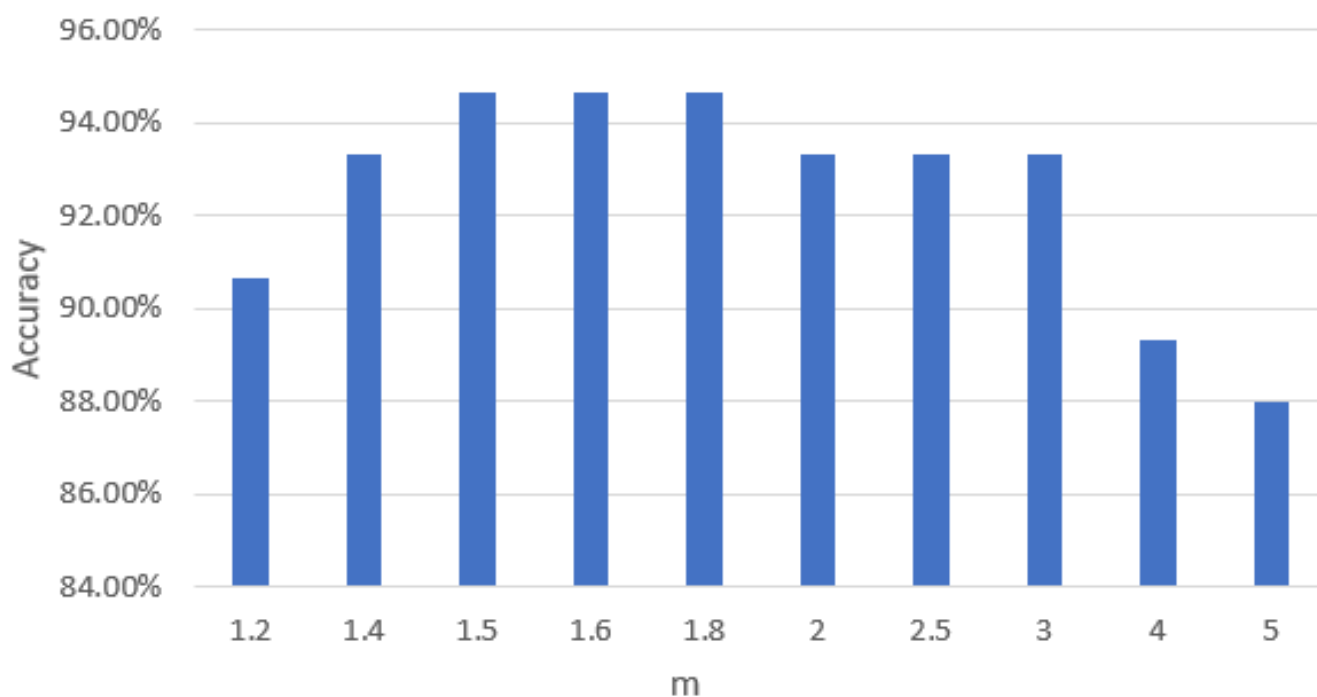


Figure 9. Classification accuracy under different weight indices.

5. Conclusions

The NIR spectra of milk samples have the question of overlap and noisy data, which brings difficulty to classifying the spectra. To increase the identification accuracy, a new fuzzy feature extraction algorithm, i.e., FiNLDA, was proposed by combining fuzzy theory with iNLDA, and it extracted the discriminant features from the NIR spectra. To classify milk brands quickly, nondestructively and effectively, the method of uniting the FiNLDA algorithm with SG filtering and PCA was designed in this study. At first, near-infrared spectra of 300 milk samples from five brands were acquired by a NIR-M-R2-type near-infrared spectrometer, and they were preprocessed by a SG-filtering algorithm. Secondly, the spectra were compressed by PCA, and extracted by LDA, iNLDA and FiNLDA, respectively. Finally, KNN was performed to identify milk brands. Compared with LDA and iNLDA, FiNLDA can accurately identify milk brands and has the highest classification accuracy. On the basis of the accurate classification of milk brands by FiNLDA, this study confirmed

the feasibility of identifying milk brands by combining the portable NIR spectrometer and FiNLDA.

Author Contributions: Conceptualization, B.W. and X.W.; methodology, X.W. and B.W.; software, B.W. and Y.F.; validation, B.W., Y.F. and X.W.; formal analysis, Y.F.; investigation, Y.F.; resources, B.W. and X.W.; data curation, B.W.; writing—original draft preparation, Y.F. and M.L.; writing—review and editing, X.W. and B.W.; visualization, M.L.; supervision, B.W. and X.W.; project administration, X.W. and B.W.; funding acquisition, Y.F. and B.W. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by Priority Academic Program Development of Jiangsu Higher Education Institutions (PAPD), the Major Natural Science Research Projects of Colleges and Universities in Anhui Province (2022AH040333), and the Undergraduate Innovation and Entrepreneurship Training Program of Jiangsu Province (202210299499X).

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: The data used to support the findings of this study can be made available by the corresponding author upon request.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Kalyankar, S.D.; Khedkar, C.; Patil, A.M.; Deosarkar, S.S. Milk: Sources and Composition. *Encycl. Food Health* **2016**, *741*–747. [[CrossRef](#)]
2. Liu, Z.; Rochfort, S. Lipidomics in milk: Recent advances and developments. *Curr. Opin. Food Sci.* **2023**, *51*, 101016. [[CrossRef](#)]
3. Medeiros, G.K.V.V.; Queiroga, R.C.R.E.; Costa, W.K.A.; Gadelha, C.A.A.; Lacerda, R.R.E.; Lacerda, J.T.; Pinto, L.S.; Braganhol, E.; Teixeira, F.C.; Barbosa, P.P.d.S.; et al. Proteomic of goat milk whey and its bacteriostatic and antitumour potential. *Int. J. Biol. Macromol.* **2018**, *113*, 116–123. [[CrossRef](#)]
4. Kandhro, F.; Kazi, T.G.; Afridi, H.I.; Baig, J.A. Compare the nutritional status of essential minerals in milk of different cattle and humans: Estimated daily intake for children. *J. Food Compos. Anal.* **2022**, *105*, 104214. [[CrossRef](#)]
5. Costa, A.; Lopez-Villalobos, N.; Sneddon, N.W.; Shaloo, L.; Franzoi, M.; De Marchi, M.; Penasa, M. Invited review: Milk lactose—Current status and future challenges in dairy cattle. *J. Dairy Sci.* **2019**, *102*, 5883–5898. [[CrossRef](#)]
6. Claeys, W.L.; Verraes, C.; Cardoen, S.; De Block, J.; Huyghebaert, A.; Raes, K.; Dewettinck, K.; Herman, L. Consumption of raw or heated milk from different species: An evaluation of the nutritional and potential health benefits. *Food Control* **2014**, *42*, 188–201. [[CrossRef](#)]
7. Palmquist, D.L.; Beaulieu, A.D.; Barbano, D.M. Feed and animal factors influencing milk fat composition. *J. Dairy Sci.* **1993**, *76*, 1753–1771. [[CrossRef](#)]
8. Gulati, A.; Galvin, N.; Hennessy, D.; McAuliffe, S.; O'Donovan, M.; McManus, J.J.; Fenelon, M.A.; Guinee, T.P. Grazing of dairy cows on pasture versus indoor feeding on total mixed ration: Effects on low-moisture part-skim Mozzarella cheese yield and quality characteristics in mid and late lactation. *J. Dairy Sci.* **2018**, *101*, 8737–8756. [[CrossRef](#)]
9. Goetsch, A.L.; Zeng, S.S.; Gipson, T.A. Factors affecting goat milk production and quality. *Small Rumin. Res.* **2011**, *101*, 55–63. [[CrossRef](#)]
10. Hayes, E.; Wallace, D.; O'Donnell, C.; Greene, D.; Hennessy, D.; O'Shea, N.; Tobin, J.T.; Fenelon, M.A. Trend analysis and prediction of seasonal changes in milk composition from a pasture-based dairy research herd. *J. Dairy Sci.* **2023**, *106*, 2326–2337. [[CrossRef](#)]
11. Schopen, G.C.B.; Heck, J.M.L.; Bovenhuis, H.; Visker, M.H.P.W.; van Valenberg, H.J.F.; van Arendonk, J.A.M. Genetic parameters for major milk proteins in Dutch Holstein-Friesians. *J. Dairy Sci.* **2009**, *92*, 1182–1191. [[CrossRef](#)]
12. Jiang, Q.; Zhang, M.; Mujumdar, A.S.; Wang, D. Non-destructive quality determination of frozen food using NIR spectroscopy-based machine learning and predictive modelling. *J. Food Eng.* **2023**, *343*, 111374. [[CrossRef](#)]
13. Folli, G.S.; Santos, L.P.; Santos, F.D.; Cunha, P.H.P.; Schaffel, I.F.; Borghi, F.T.; Barros, I.H.A.S.; Pires, A.A.; Ribeiro, A.V.F.N.; Romão, W.; et al. Food analysis by portable NIR spectrometer. *Food Chem. Adv.* **2022**, *1*, 100074. [[CrossRef](#)]
14. Alaoui Mansouri, M.; Ziemons, E.; Sacré, P.Y.; Kharbach, M.; Barra, I.; Cherrah, Y.; Hubert, P.; Marini, R.D.; Bouklouze, A. Classification of polymorphic forms of fluconazole in pharmaceuticals by FT-IR and FT-NIR spectroscopy. *J. Pharmaceut. Biomed.* **2021**, *196*, 113922. [[CrossRef](#)]
15. Larrechi, M.S.; Callao, M.P. Strategy for introducing NIR spectroscopy and multivariate calibration techniques in industry. *TrAC Trends Anal. Chem.* **2003**, *22*, 634–640. [[CrossRef](#)]
16. Blanco, M.; Coello, J.; Iturriaga, H.; Maspocho, S.; Rovira, E. Determination of water in ferrous lactate by near infrared reflectance spectroscopy with a fibre-optic probe. *J. Pharmaceut. Biomed.* **1997**, *16*, 255–262. [[CrossRef](#)]

17. Zhang, Y.; Liao, T.; Wang, G.; Xu, J.; Ren, F.; Zhang, H. An ultrasensitive NIR-IIa' fluorescence-based multiplex immunochromatographic strip test platform for antibiotic residues detection in milk samples. *J. Adv. Res.* **2022**, *50*, 25–34. [[CrossRef](#)] [[PubMed](#)]
18. Said, M.; Wahba, A.; Khalil, D. Semi-supervised deep learning framework for milk analysis using NIR spectrometers. *Chemometr. Intell. Lab.* **2022**, *228*, 104619. [[CrossRef](#)]
19. Nagy, M.M.; Wang, S.; Farag, M.A. Quality analysis and authentication of nutraceuticals using near IR (NIR) spectroscopy: A comprehensive review of novel trends and applications. *Trends Food Sci. Technol.* **2022**, *123*, 290–309. [[CrossRef](#)]
20. Zhang, J.; Guo, Z.; Wang, S.H.; Yue, M.H.; Zhang, S.S.; Peng, H.H.; Yin, X.; Du, J.; Ma, C.Y. Comparison of methods for water content in rice by portable near-infrared and visible light spectrometers. *Spectrosc. Spectr. Anal.* **2023**, *43*, 2059–2066.
21. Chen, L.F.; Liao, H.Y.M.; Ko, M.T.; Lin, J.C.; Yu, G.J. A new LDA-based face recognition system which can solve the small sample size problem. *Pattern Recogn.* **2000**, *33*, 1713–1726. [[CrossRef](#)]
22. Ye, J.P.; Tao, X. Computational and theoretical analysis of null space and orthogonal linear discriminant analysis. *J. Mach. Learn. Res.* **2006**, *7*, 1183–1204.
23. Shen, Y.; Wu, X.; Wu, B.; Tan, Y.; Liu, J. Qualitative analysis of lambda-cyhalothrin on Chinese cabbage using mid-infrared spectroscopy combined with fuzzy feature extraction algorithms. *Agriculture* **2021**, *11*, 275. [[CrossRef](#)]
24. Fábíán, G.; Generalized, S.G. Filter for smoothing triangular meshes. *Comput. Aided Geom. Des.* **2023**, *100*, 102167. [[CrossRef](#)]
25. Schafer, R.W. What is a Savitzky-Golay Filter. *IEEE Signal Proc. Mag.* **2011**, *28*, 111–117. [[CrossRef](#)]
26. Zhang, T.; Wu, X.; Wu, B.; Dai, C.; Fu, H. Rapid authentication of the geographical origin of milk using portable near-infrared spectrometer and fuzzy uncorrelated discriminant transformation. *J. Food Process Eng.* **2022**, *45*, e14040. [[CrossRef](#)]
27. Li, B.; Wei, Y.; Duan, H.; Xi, L.; Wu, X. Discrimination of the geographical origin of *Codonopsis pilosula* using near infrared diffuse reflection spectroscopy coupled with random forests and k-nearest neighbor methods. *Vib. Spectrosc.* **2012**, *62*, 17–22. [[CrossRef](#)]
28. An, M.H.; Cao, C.M.; Wang, S.S.; Zhang, X.C.; Ding, W.Y. Non-destructive identification of moldy walnut based on NIR. *J. Food Compos. Anal.* **2023**, *121*, 105407. [[CrossRef](#)]
29. Bezdek, J.C. *Pattern Recognition with Fuzzy Objective Function Algorithms*; Plenum Press: New York, NY, USA, 1981.

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.