Current Research in Food Science 9 (2024) 100929

Contents lists available at ScienceDirect

Current Research in Food Science

journal homepage: www.sciencedirect.com/journal/current-research-in-food-science

Skin hyperspectral imaging and machine learning to accurately predict the muscular poly-unsaturated fatty acids contents in fish

Yi-Ming Cao [a], Yan Zhang [a], Qi Wang [a], Ran Zhao [a], Mingxi Hou [a], Shuang-Ting Yu [a,b], Kai-Kuo Wang [c], Ying-Jie Chen [c], Xiao-Qing Sun [a], Shijing Liu [d], Jiong-Tang Li [a,*]

[a] Key Laboratory of Aquatic Genomics, Ministry of Agriculture, Beijing Key Laboratory of Fishery Biotechnology, Chinese Academy of Fishery Sciences, Beijing, 100141, China
[b] Chinese Academy of Agricultural Sciences, Beijing, 100141, China
[c] National Demonstration Center for Experimental Fisheries Science Education, Shanghai Ocean University, Shanghai, 201306, China
[d] Fishery Machinery and Instrument Research Institute, Chinese Academy of Fishery Sciences, Shanghai, 200092, China

ARTICLE INFO

Handling Editor: Dr. Xing Chen

Keywords:
PUFAs
DHA
EPA
Hyperspectral imaging
Machine learning

ABSTRACT

The polyunsaturated fatty acids (PUFAs), particularly eicosapentaenoic acid (EPA) and docosahexaenoic acid (DHA), are critical determinants of the nutritional quality of fish. To rapidly and non-destructively determine the muscular PUFAs in living fish, an accuracy technique is urgently needed. In this study, we combined skin hyperspectral imaging (HSI) and machine learning (ML) methods to assess the muscular PUFAs contents of common carp. Hyperspectral images of the live fish skin were acquired in the 400–1000 nm spectral range. The spectral data were preprocessed using Savitzky–Golay (SG), multivariate scattering correction (MSC), and standard normal variable (SNV) methods, respectively. The competitive adaptive reweighted sampling (CARS) method was applied to extract the optimal wavelengths. With the skin spectra of fish, five ML methods, including the extreme learning machine (ELM), random forest (RF), radial basis function (RBF), back propagation (BP), and least squares support vector machine (LS-SVM) methods, were used to predict the PUFAs and EPA + DHA contents. With the spectral data processed with the SG, the RBF model achieved outstanding performance in predicting the EPA + DHA and PUFAs contents, yielding coefficients of determination (R_P^2) of 0.9914 and 0.9914, root mean square error (RMSE) of 0.3352 and 0.3346, and mean absolute error (MAE) of 0.2659 and 0.2660, respectively. Finally, the visualization distribution maps under the optimal model would facilitate the direct determination of the fillet PUFAs and EPA + DHA contents. The combination of skin HSI and the optimal ML method would be promising to rapidly select living fish having high muscular PUFAs contents.

1. Introduction

Fatty acids (FAs) participate in many essential physiological functions of organisms (De Carvalho, C., and Caramujo, M., 2018). Among FAs, polyunsaturated fatty acids (PUFAs), such as eicosapentaenoic acid (EPA, C20:5n-3) and docosahexaenoic acid (DHA, C22:6n-3), are essential for the growth, metabolism and physiological processes of humans (Zhang et al., 2019; Jamshidi et al., 2020; Rohman et al., 2021). For humans, they play an important role in neuroimmunity, and neuroprotection and effectively contribute to the prevention of cardiovascular disease, reducing blood cholesterol levels and decreasing the probability of blood clots (Petel et al., 2007; Djuricic and Calder, 2021). Fish are among the main sources of PUFAs for humans. Therefore, it is necessary to measure the contents of PUFAs in fish fillets. At present, gas chromatography (GC) is the conventional method for assessing the FAs contents (Ecker et al., 2012; Chiu and Kuo, 2020). However, GC analysis involves cumbersome processing steps, including extensive sample preparation, preparation of organic solvents, and time-consuming assay procedures (Rohman et al., 2021; Xia et al., 2019). Therefore, it is crucial to develop a rapid and noninvasive method to assess the PUFAs contents in fish fillets.

Recently optical techniques including near-infrared spectroscopy

* Corresponding author.
E-mail addresses: caoyiming@cafs.ac.cn (Y.-M. Cao), zhangy@cafs.ac.cn (Y. Zhang), wangqi@cafs.ac.cn (Q. Wang), zhaoran@cafs.ac.cn (R. Zhao), houmingxi@cafs.ac.cn (M. Hou), styuwork@163.com (S.-T. Yu), 18631836881@163.com (K.-K. Wang), cyjttkl@163.com (Y.-J. Chen), sunxiaoqing@cafs.ac.cn (X.-Q. Sun), liushijing@fmiri.ac.cn (S. Liu), lijt@cafs.ac.cn (J.-T. Li).

https://doi.org/10.1016/j.crfs.2024.100929
Received 28 July 2024; Received in revised form 14 November 2024; Accepted 16 November 2024
Available online 18 November 2024

(NIRS) and Raman spectroscopy have been widely used for identifying food components (Abbas et al., 2012). NIRS, a rapid non-destructive method with minimal sample preparation, is suitable for assessing food quality (Chapman et al., 2020; Lintvedt et al., 2023). However, this approach is limited to a single point in the preselected areas, the result of which cannot represent the whole sample. Raman spectroscopy faces challenges from weak Raman scattering signals (Afseth et al., 2022; Xu et al., 2015). Raman spectroscopy is difficult to measure substances of low content, which limits its wide application in the food industry (Petersen et al., 2021). Additionally, its small focal volume meant that the measurements may not represent the entire sample (Lintvedt et al., 2023). Furthermore, the large equipment is not suitable for the field detection. Hyperspectral imaging (HSI) integrates spectral and image data, facilitating non-destructive acquisition of two-dimensional images and one-dimensional spectral data from samples (Berhe et al., 2016; Ma et al., 2018;Zhang and Abdulla, 2022). HSI has been used as a non-destructive technique to assess meat nutrition and freshness (Cheng et al., 2019, 2023; Wang et al., 2020; Lintvedt et al., 2023). However, these studies focused on processed meat with the HSI technique. Whether this technique could be used to assess the nutrition values of living animal muscle has not been studied. If feasible, this technique would assist in selecting living animals having high muscular nutrition values in advance before the following process.

Skin spectral images provide information on the spatial distribution and concentration of biological components (He and Wang, 2019). In humans, skin HSI data have been used to evaluate the body's pathological or physiological status (Yudovsky et al., 2011; Paul et al., 2015; Ross et al., 2014; Zherebtsov et al., 2019). Although the skin HSI data have not been employed in assessing the muscular nutrition values of living animals, it provides an opportunity to generate a non-destructive method combining the skin HSI data and the machine learning (ML) method to predict the muscular PUFA contents of living animals. In this study, we obtained the spectral information of fish skin and established the optimal relationship between skin information and muscle PUFAs contents using five machine learning (ML) methods with high prediction accuracy. Our findings could facilitate rapid, efficient, and non-destructive detection of PUFAs contents in living fish.

2. Materials and methods

2.1. Sample collection and preparation

We collected 400 healthy and uninjured live common carp from the Experimental Fish Farm of the Chinese Academy of Fisheries Sciences (Fangshan, Beijing). The samples were approximately one year old with an average weight of approximately 607.20 g. For each sample, the skin and corresponding muscle were collected from a region different from the previous dorsal region (Cao et al., 2023). After removing the water, the skin of each live fish was scanned using the HSI system. Then we scraped the skin and incised the corresponding muscle to determine the PUFAs contents. The sample preparation details are provided in our previous study (Cao et al., 2023). The key steps in the process are summarized in Fig. 1.

2.2. Determination of FAs contents

A gas chromatography (GC) system was employed to determine the PUFAs contents in the muscle of common carp. Lipids were extracted from common carp muscle, and the content of each muscle sample was determined by the gravimetric method (Sushchik et al., 2007). Fatty acid methyl esters (FAMEs) were obtained from lipids (Ma et al., 2018). An Agilent 7890 A GC system (Agilent Technologies, Wilmington, DE, USA) equipped with a flame ionization detector and an Agilent DB-23 capillary column (Agilent Technologies, Wilmington, DE, USA) was utilized for the detection of FAMEs (Zhang et al., 2021). The initial temperature of the GC was set at 60 °C for 1 min, then increased to 150 °C at a rate of 10 °C/min, followed by 200 °C at 15 °C/min, and finally to 230 °C at 30 °C/min. The final temperature was maintained for 5 min. The injection and detection temperatures were maintained at 300 °C. The carrier gas pressure was set at 40 kPa, with nitrogen used as the carrier gas. The type of each FA was determined by comparing the retention time of the samples with the peaks of a Supelco 37-component FAME standard combination (Nu-Chek Prep, Inc., Elysian, USA). The relative content of each FA was calculated using area normalization (Sadeghin et al., 2018). The percentage of each PUFA was calculated as (its area/total area of all FAs) × 100. The absolute level of each FA was multiplied by the total lipid content of the sample to calculate the absolute content of each FA. The absolute content of total PUFA was equal to the absolute content sum of each detected PUFA. The absolute content of EPA + DHA was the absolute sum of the contents of C20:5n-3 and C22:6n-3.

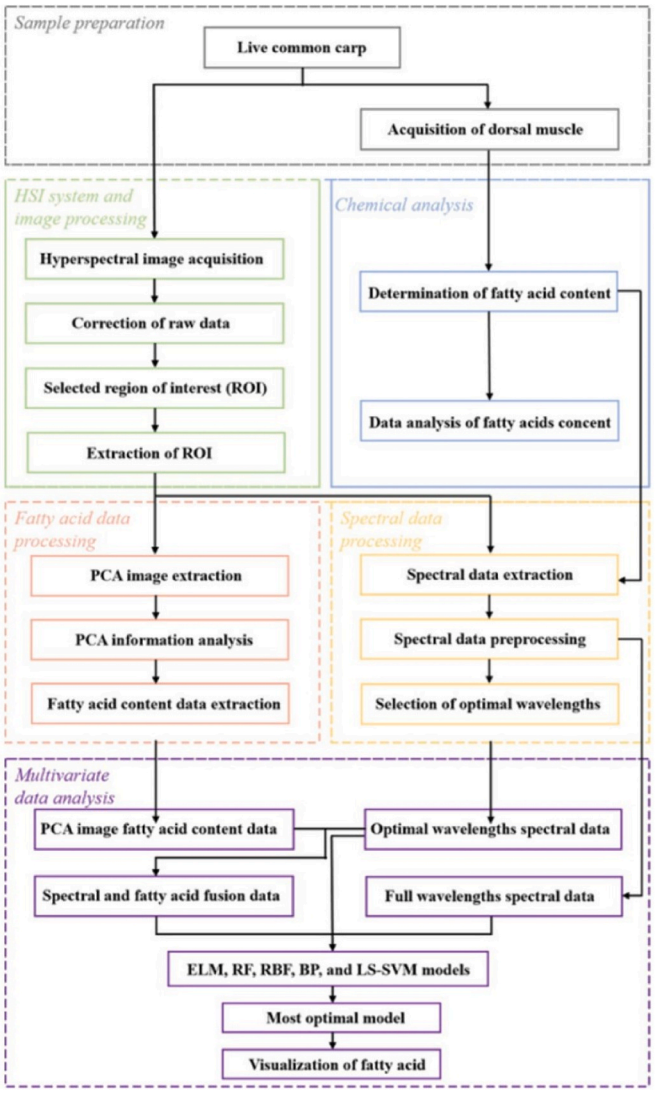


Fig. 1. Flow chart of the main steps of data collection, analysis, and modeling.

2.3. Image capture using a hyperspectral imaging system

An HSI system consisting of a hyperspectral imager (FigSpec Hyperspectral Camera FS-13, FigSpec Technology, Zhejiang) with a spectral range of 400–1000 nm, a mobile platform, six 50 W fibre-optic halogen lamps, a dark box (50 cm × 40 cm × 30 cm) for reducing ambient light, and a LENOVO computer equipped with an image acquisition card. The spectral resolution was 2.5 nm, the slit width was

25 μm, and the transmission efficiency was >60%. The scaled common carp was placed side up on a moving platform to obtain a hyperspectral image of the skin. The original hyperspectral image was corrected to eliminate the effects of dark current effects and uneven illumination from the hyperspectral camera. The detailed image-scanning process was described in our previous study (Cao et al., 2023). Briefly, the imaging speed was 128 Hz over the full wavelength range and the scanning speed was 30 rows/second. The distance and intensity of the illumination source were adjusted to ensure the clarity of the hyperspectral image before acquiring the hyperspectral image of each sample.

2.4. Extraction of the reflectance spectra of the region of interest

Following Kamruzzaman et al.'s method (Kamruzzaman et al., 2016), the regions of interest (ROIs) in the hyperspectral images were identified and selected. A region of 200 pixels × 200 pixels with the exposed skin at each corrected hyperspectral image was selected using the rectangle tool in ENVI v5.3 software (Exelis Visual Information Solutions, Inc., Boulder, CO, USA). The background was eliminated by subtracting the sample's high reflection band from the sample's low reflection band, and then masking. The ROIs of the sample image was extracted by setting a threshold (0.075). The reflectance spectra of all the pixels were averaged as the spectral data of the ROI.

2.5. Spectral data analysis

2.5.1. Preprocessing the spectral data

The 400 samples were randomly divided into two groups using the Kennard–Stone (KS) method (Sun et al., 2021). The first set consisted of 320 samples for the construction of the calibration model (calibration set) and the other samples were used to validate the accuracies of the models (prediction set). Preprocessing minimizes the adverse effects caused by environmental factors such as light scattering. To compare the effects of the preprocessing methods on the modeling and select the optimal method, we employed three methods, including Savitzk–Golay (SG) smoothing, multiplicative scatter correction (MSC), and standard normal variate (SNV) (Ravikanth et al., 2017).

2.5.2. Selecting the optimal wavelength

The HSI provides a large amount of spatial and spectral information about the sample. However, redundant data at certain wavelengths might slow the modeling. Therefore, it is necessary to select the optimal wavelengths to represent the full wavelengths and reduce the data dimensionality to improve computational efficiency (Yu et al., 2021). Competitive adaptive reweighted sampling (CARS) is effective in reducing the dimensionality of wavelength information and selecting key variables to improve the predictive performance of the model (Li et al., 2009). CARS selects a subset of variables from Monte Carlo (MC) sampling using iterative and competitive means (Cheng et al., 2020). Herein, the MC sample size was set to 50. Using multiple subsets of variables coupled with different sample sizes, the cross-validation method was used to calculate the root means square error of the cross-validation (RMSECV) for each subset. If one subset had the smallest RMSECV, these variables in this subset were the optimal combination. The fold of cross-validation was set as five.

2.5.3. Predicting the muscle with five machine learning methods

Five ML methods, the extreme learning machine (ELM), random forest (RF), radial basis function (RBF), back propagation (BP), and least squares support vector machine (LS-SVM) methods, were used to develop a calibration model to predict the muscular EPA + DHA and PUFAs contents of common carp. The ELM is an ML method based on a feed-forward neural network (Chen et al., 2014). The ELM consists of an input layer, a hidden layer, and an output layer. The input layer has N neurons corresponding to N input variables. The hidden layer has one neuron, and the output layer has M neurons corresponding to M output variables. The weights from the input layer to the hidden layer and the deviation of the hidden layer can be randomly initialized without iterative correction so that some computations can be reduced. The weights from the hidden layer to the output layer can be obtained by solving the matrix equations.

The RF is a supervised ML algorithm that uses a bootstrap statistical resampling technique to create a collection of decision trees (Saha et al., 2021). The training dataset of each tree in the forest is sampled with replacement, and a random subset of features is used at each decision split. The number of labels classified by all decision trees directly determines the final prediction of a forest. The tree number and leaf nodes of the model were set to 100 and 2, respectively.

The RBF is an artificial neural network that performs supervised ML (Saha et al., 2021; Khan et al., 2022). It is a more complex process to analyze a large amount of data than simple linear classifiers such as logistic regression and linear discriminant analysis. The radial basis function of the model was set to expand at a rate of 1000.

The BP is a classical feed-forward multilayer (i.e., an input layer, one or more hidden layers, and an output layer) and is a unidirectional connection (weight) from input to output (Saha et al., 2021). The number of hidden layer nodes, the learning rate factor, the momentum factor, the initial weight, and the number of iterations were set as 5, 0.1, 0.1, 0.3, and 1000, respectively. The scale function was set to the "tanh" function.

The LS-SVM uses the radial basis kernel function (RBF), which is a nonlinear function that reduces the complexity of the training process (Saha et al., 2021). The regularization parameter gamma (γ) and the kernel parameter (σ^2) represent the width of the RBF kernel and can reduce the complexity. The γ and σ^2 were set to 10 and 2, respectively.

2.5.4. Model building and performance evaluation

The performance of the prediction model was evaluated with four statistical parameters, including the calibration set coefficient of determination (R_C^2), prediction set coefficient of determination (R_P^2), root mean square error (RMSE), and mean absolute error (MAE). A prediction model has excellent performance when R^2 is high and RMSE and MAE are low (Zheng et al., 2019; Aiyelokun and Agbede, 2021). The R^2, RMSE, and MAE are calculated as shown in equations (1)–(3):

$$R^2 = \sum_{i=1}^{n}(x_i-\overline{x})^2(y_i-\overline{y})^2 \Big/ \sum_{i=1}^{n}(x_i-\overline{x})^2\sum_{i=1}^{n}(y_i-\overline{y})^2 \quad (1)$$

$$RMSE = \sqrt{1\Big/k\sum_{i=1}^{n}(x_i-y_i)^2} \quad (2)$$

$$MAE = \sum_{i=1}^{n}|x_i-y_i|\Big/k \quad (3)$$

where n represents the number of samples, x_i and y_i represents the estimated values and measured values, respectively, and $\overline{x}$ $\overline{y}$ represents the average estimated values and measured values, respectively.

2.6. Visualization of the DHA + EPA and PUFAs contents

The hyperspectral data of the samples were processed in a pixel-wise format by inputting the spectra of each pixel into the best calibration model. Subsequently, the predicted value for each pixel was visualized through a distribution plot, representing the actual values concerning their respective indices in the original hyperspectral image. This approach allowed us to determine the distributions of PUFAs and EPA + DHA in the muscle of live common carp. The visualized distribution maps are depicted using linear color bars.

2.7. Statistical analysis

The obtained data are expressed as mean ± standard deviation. Statistical analyses were performed using SPSS version 25 software (IBM Corp, Armonk, NY). Student's t-test was used in comparing the means of two sets of continuous variables, assuming that the data were normally distributed. $P < 0.05$ was considered statistically significant.

3. Results

3.1. Variable contents of PUFAs and EPA + DHA in the common carp muscle

The PUFAs and EPA + DHA contents of the dorsal muscle of 400 common carp individuals were obtained (Table S1). The EPA + DHA content in the calibration samples ranged from 0 to 2.38 mg/g, with a standard deviation of 0.46 (Fig. 2a), indicating that there was variability among the samples. The predicted samples exhibited a comparable range of EPA + DHA contents, ranging from 0 to 2.03 mg/g with a standard deviation of 0.44 (Fig. 2a). For PUFAs, the contents in the calibration samples varied from 1.15 to 17.91 mg/g, while those in the prediction set samples ranged from 1.36 to 18.89 mg/g (Fig. 2b). The means and standard deviations of the calibration and prediction sets were similar for both EPA + DHA and PUFAs, and the distributions of these two indicators were not significantly different (independent samples *t*-test, P values of 0.979 and 0.651, respectively; Table 1). Hence, for these two parameters, the prediction set was applied to validate the accuracy of the prediction model generated from the calibration set.

3.2. Diverse preprocessed skin spectral characteristics obtained by different methods

The spectral data of 400 live common carp skin samples in the wavelength range of 400–1000 nm were documented in Table S2. Although we selected different skin regions of the same samples from our previous study, the raw spectra of live common carp skin exhibited similar trends in these two studies (Fig. 3a). Furthermore, the reflectance values from fish skin had a similar distribution to the muscular reflectance data (Cheng et al., 2019). Since fish muscular reflectance data have been widely used to predict muscular nutrient contents (Cheng and Sun, 2017), it is possible that skin spectra could be used to predict muscular nutrient contents in live fish.

The SG method smoothed spectra and eliminated the effect of noise generated by relevant experimental factors. However, in general, the spectral trends remained similar to those of the raw spectra (Fig. 3b). In particular, the reflectivity between 500 and 900 nm ranged from 0 to 60 with a similar distribution in the raw spectra. In contrast, the reflectivity of the fish skin spectra at 500–900 nm were compressed into the range between 5 and 25 (Fig. 3c). Similarly, after applying the MSC method, the reflectivity between 500 and 900 nm was further reduced to a range from −2 to 1.5 (Fig. 3d). These data revealed that the reflectivity between 500 and 900 nm was greatly affected by different preprocessing methods.

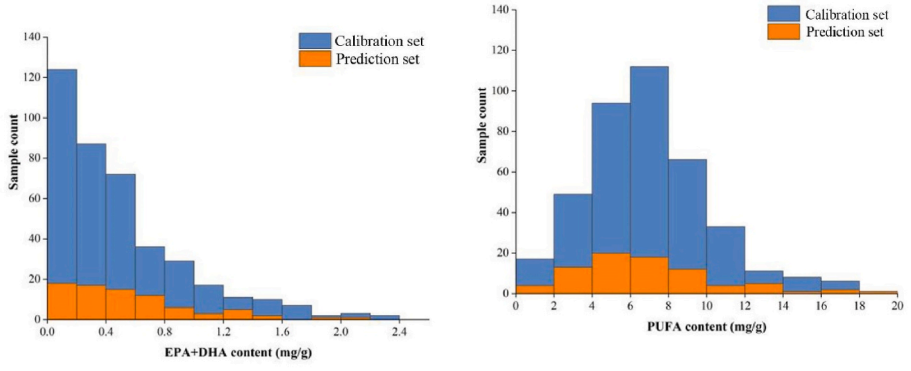
3.3. Common characteristic spectra of PUFA contents and EPA + DHA contents using the CARS method

As the number of MC samples increased, using either the raw spectra or the preprocessed spectra with SG, MSC, and SNV, the number of characteristic wavelengths rapidly decreased in the initial stage (the first 10 sampling processes) and subsequently decreased slowly, revealing the two phases of quick selection and fine selection (Fig. 4a–S1a, S2a, and S3a).

The five-fold RMSECV value distributions of the raw spectra gradually decreased until the 22nd sampling. Then the RMSECV values were stable along with more sampling (Fig. 4b). The minimum value of the RMSECV was obtained when the sampling time was run to 48. However, the RMSECV value distributions using the preprocessing data from three methods were different from the above distribution. The five-fold RMSECV values of the SG preprocessed spectra showed a decrease at first and then an increase (Fig. S1b). The RMSECV values of the spectra after MSC and SNV treatments gradually increased with the sampling runs (Figs. S2b and S3b). With these three sets of preprocessed data, the minimum values of RMSECV were obtained when the sampling time was 27, 11, and 12, respectively. The change process of the regression coefficients paths with an increase in sampling times was described in Fig. 4c–S1c, S2c, and **S3c**.

The optimal wavelengths selected by the CARS method for EPA + DHA and PUFAs prediction are summarized in Table 2. The SG-CARS screened 14 optimal wavelengths including 395, 411, 427, 655, 711, 968, 987, 989, 1000, 1004, 1028, 1030, 1033, and 1035 nm for the prediction of the EPA + DHA content. For the MSC-CARS, four optimal variables (395, 405, 993, and 1013 nm) were selected. The SNV-CARS chose 20 optimal wavelengths (393, 395, 398, 414, 421, 499, 501, 504, 506, 508, 510, 513, 515, 957, 976, 983, 985, 1026, 1028, and 1033 nm) for determining EPA + DHA content. Intriguingly, two spectral regions of 390–420 nm and 950–1000 nm were selected by all three methods, suggesting that these two regions might reflect the spectral characteristics of the EPA + DHA content.

To estimate the PUFAs content, 11 optimal wavelengths (936, 968, 1013, 1017, 1019, 1024, 1026, 1028, 1030, 1032, and 1035 nm) were selected by the SG-CARS method. The MSC-CARS method revealed 26 optimal wavelengths (418, 421, 423, 432, 446, 477, 479, 482, 495, 519,



(a) (b)

Fig. 2. Comparison of sample distributions of EPA + DHA (a) and PUFAs (b) contents between the calibration and prediction sets. The numbers on the Y-axis represent the value of the sample count.

Table 1
Descriptive statistics of absolute EPA + DHA and PUFA contents (mg/g) of common carp muscle.

Fatty acids	Calibration set					Prediction set					*t*-test
	N	Mean	Max	Min	SD	N	Mean	Max	Min	SD	P
EPA + DHA	320	0.47	2.38	0	0.46	80	0.56	2.03	0	0.44	0.979
PUFA	320	7.10	17.91	1.15	4.10	80	7.01	18.89	1.36	3.63	0.651

Note: N: number of samples; SD: standard deviation.

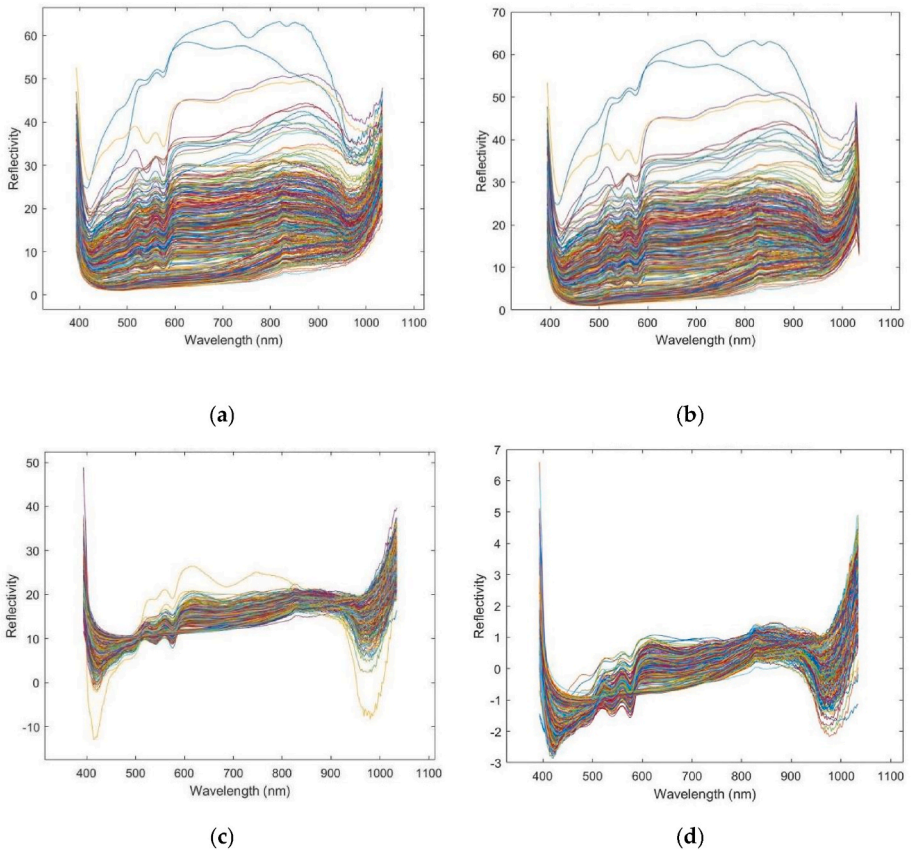


Fig. 3. Mean spectra of common carp obtained with different preprocessing methods: (a) original mean spectra; (b) mean spectra preprocessed by SG; (c) mean spectra preprocessed by MSC; (d) mean spectra preprocessed by SNV. The different colored curves represent the average spectra of 400 samples. The numbers on the X-axis represent the Wavelength. The numbers on the Y-axis represent the value of the reflectivity.

528, 550, 561, 584, 591, 655, 657, 661, 674, 726, 757, 974, 1017, and 1024 nm). The optimal wavelengths chosen by the SNV-CARS method were the fewest with only ten wavelengths (951, 972, 983, 993, 998, 1002, 1009, 1026, 1030, and 1032 nm). Intriguingly, reflectance values ranging from 950 nm to 1000 nm were selected by all three methods, suggesting that the data in this range might reflect the spectral characteristics of PUFAs content.

The reflectance in the 950–1000 nm region corresponds to C-H and O-H bonds in FAs (Wu and Sun, 2013; Kamper et al., 2020). Hence, the 950–1000 nm spectra could be used to predict the FAs contents (Tahmasbian et al., 2021). Furthermore, the spectra ranging from 916 nm to 1000 nm were applied to predict the contents of EPA and DHA in the fillets of salmon and grass carp (Cheng et al., 2019). Our optimal wavelengths were consistent with those used in previous studies. Although we used three different preprocessing methods, the CARS method identified the common characteristic spectra of the PUFAs and EPA + DHA contents.

3.4. High prediction accuracies of the contents of PUFAs and EPA + DHA

With the selected optimal wavelengths, we first estimated the prediction accuracy of different modeling methods for the EPA + DHA content (Table 3). With the SG data, the LS-SVM and RBF models demonstrated superior performance, achieving R_P^2 values of 0.9913 and 0.9914, respectively, with RMSE values of 0.3398 and 0.3352 for the prediction set. Using the MSC data, the LS-SVM and BP models performed relatively well, with R_P^2 values of 0.7526 and 0.7413, respectively. Using the SNV preprocessed data, only the ELM prediction model achieved an R_P^2 greater than 0.88. The results suggested that even when using the same preprocessing method, the prediction accuracies of the modeling methods were very different.

The effects of various preprocessing models on the prediction accuracies were also compared with those of the same modeling method (Table 3). For the LS-SVM and RBF models, the SG data exhibited better prediction performance (R_P^2 values of 0.9913 and 0.9914, respectively) than did the other two preprocessed datasets (MSC: R_P^2 values of 0.7526 and 0.7395, respectively; SNV: R_P^2 values of 0.6220 and 0.5086, respectively). For the ELM model, the SNV model ($R_P^2 = 0.7312$) outperformed the SG ($R_P^2 = 0.5104$) and MSC ($R_P^2 = 0.6401$) models. The spectral data preprocessed by MSC fit well with the BP model ($R_P^2 = 0.7413$), which yielded superior predictions compared to those obtained using SG ($R_P^2 = 0.6626$) and SNV ($R_P^2 = 0.5517$). The results suggested

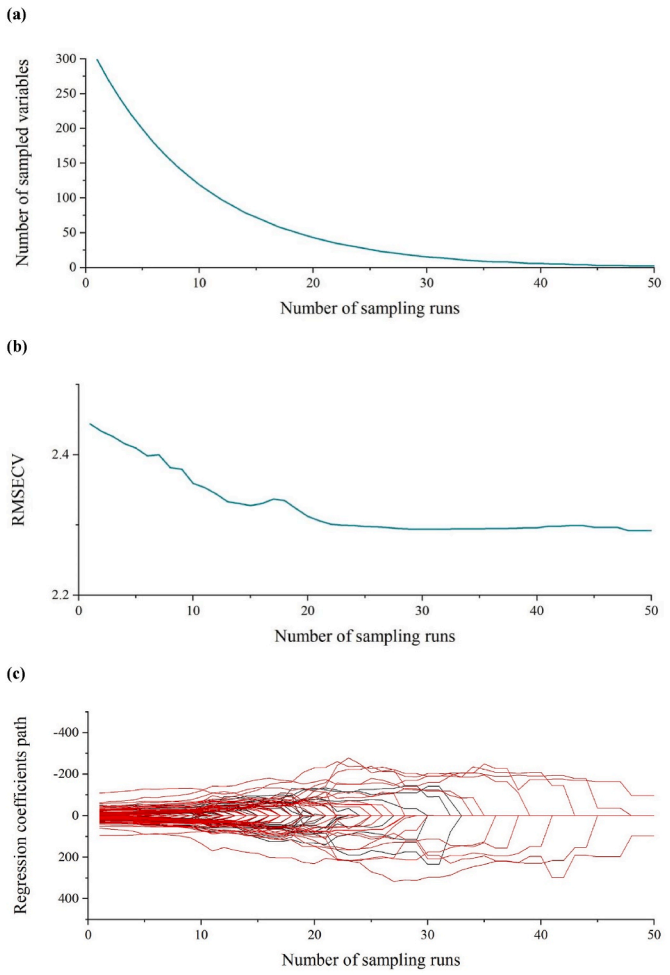


Fig. 4. CARS selects the raw spectral variables. (a) The trend in the amount of features sampled; (b) RMSECVs with different MC sampling runs; (c) coefficient paths as a function of sampling number. The numbers on the X-axis represent the number of sampling runs. The numbers on the Y-axis represent the number of sampled variables, RMSECV, and regression coefficients path, respectively.

Table 2
The optimal wavelengths for EPA + DHA and PUFAs contents prediction in the different preprocessing methods selected by the CARS method.

Index	Preprocessing method	Wavelength number	Feature wavelength
EPA + DHA	SG	14	395, 411,427, 655, 711, 968, 987, 989, 1000, 1004, 1028, 1030, 1033, 1035
	MSC	4	395, 405, 993, 1013
	SNV	20	393, 395, 398, 414, 421, 499, 501, 504, 506, 508, 510, 513, 515, 957, 976, 983, 985, 1026, 1028, 1033
PUFAs	SG	11	936, 968, 1013, 1017, 1019, 1024, 1026, 1028, 1030, 1032, 1035
	MSC	26	418, 421, 423, 432, 446, 477, 479, 482, 495, 519, 528, 550, 561, 584, 591, 655, 657, 661, 674, 726, 757, 974, 1017, 1024
	SNV	10	951, 972, 983, 993, 998, 1002, 1009, 1026, 1030, 1032

Note: SG: savitzk–Golay. MSC: multiplicative scatter correction. SNV: standard normal variate.

that even when using the same modeling method, the prediction accuracies of the preprocessing methods were much different. The SG-CARS-LV-SVM combination (R_P^2 value of 0.9913; Fig. 5a) and the SG-CARS-RBF combination (R_P^2 value of 0.9914; Fig. 5b) had the highest accuracies in predicting the muscle EPA + DHA contents using the skin optimal spectra.

We also estimated the prediction accuracy of different modeling methods on the PUFAs content with the same preprocessing method (Table 4). With the SG method, the LS-SVM and RBF models exhibited the best prediction performances, with R_P^2 values of 0.9910 and 0.9914, respectively. Both RMSE values were less than 0.346. With the MSC method, the LS-SVM and BP models exhibited superior predictive performance, achieving R_P^2 values of 0.7387 and 0.7583, respectively. Conversely, the highest prediction accuracy (LS-SVM model with an R_P^2 value of 0.6408) of the five prediction models using the SNV processed data was lower than the best accuracies using the processed data from the former two methods. This result demonstrated that the modeling method also affected the prediction accuracy of PUFAs content.

Finally, we examined whether the preprocessing methods affected the accuracies of the prediction models on the PUFAs contents (Table 4). With the MSC processed data, the ELM prediction model demonstrated superior performance ($R_P^2 = 0.6604$) compared to the SG ($R_P^2 = 0.4588$) and SNV ($R_P^2 = 0.4541$) models. When employing the SG data, the LS-SVM, RF, RBF, and BP prediction models achieved higher R_P^2 values (0.9910, 0.6909, 0.9914, and 0.9782, respectively) than those data obtained with the MSC method (0.7387, 0.0329, 0.0739, and 0.7583, respectively) and SNV method (0.6408, 0.5751, 0.5708, and 0.5108, respectively). The results confirmed the preprocessing methods as another key factor affecting the prediction accuracies of PUFAs content. The SG-CARS-LS-SVM, SG-CARS-RBF, and SG-CARS-BP exhibited higher R_P^2 values (0.9910, 09914, and 0.9782; Fig. 5c, d, and 5e) and lower RMSE values (0.3452, 0.3352, and 0.5326, respectively). Intriguingly, the SG-CARS-RBF combination had the best prediction accuracy for the contents of both EPA + DHA and PUFAs with the highest R_P^2 values (0.9914 and 0.9914).

3.5. Visualization for determination of fillet PUFAs content and EPA + DHA content

The RBF model was selected as the best prediction model for these two nutrient traits. The predicted values were used to visualize the EPA + DHA and PUFAs contents. We selected the samples with EPA + DHA contents at the lower quartile, median, upper quartile, and the highest quartile for visualization. In the distribution map, along with the increasing EPA + DHA contents, the number of red areas representing the high EPA + DHA content was rising (Fig. 6a). We also selected the samples for the PUFAs content visualization using the same strategy. The increasing content trend was observed to be represented with more red areas (Fig. 6b).

Using the distribution maps, we found that the muscles were split into multiple irregular blocks with intervals of low nutrient contents. The contents of the blocks determined the overall level of the ROI. Furthermore, the contents of the blocks in one ROI were greatly different. Visualization could be useful for directly determining the nutrient contents and facilitating the selection of fish with high nutrient levels.

4. Discussion

Our study revealed that the fish skin spectra detected with the HSI system could be utilized to predict the muscular contents of PUFAs and EPA + DHA with very high accuracy. Traditional GC methods are less representative because they take a long time and analyze only one sample (Rohman et al., 2021; Xia et al., 2019). The traditional GC method makes it difficult to measure the PUFAs contents of the entire samples. Additionally, the GC method is time-consuming and destructive for sample preparation (Rohman et al., 2021; Xia et al., 2019). Our

Table 3
Performance of modeling models to predict EPA + DHA contents using different preprocessing methods and characteristic wavelength screening methods.

Preprocessing Method	Band Selection	Modeling Method	Calibration			Prediction		
			R_C^2	RMSE	MAE	R_P^2	RMSE	MAE
SG	CARS	ELM	0.6463	1.6606	0.8311	0.5104	2.0423	1.4061
		LS-SVM	**0.9934**	**0.3337**	**0.2692**	**0.9913**	**0.3398**	**0.1155**
		RF	0.8185	1.7442	0.8434	0.7038	1.9537	1.3497
		RBF	**0.9934**	**0.3326**	**0.2704**	**0.9914**	**0.3352**	**0.2659**
		BP	0.6834	1.8302	1.5312	0.6626	0.2560	0.2035
MSC	CARS	ELM	0.6959	2.6919	1.5749	0.6401	1.6974	0.8313
		LS-SVM	0.7580	2.0983	1.6008	0.7526	1.9720	1.5583
		RF	0.5654	2.6994	1.5699	0.0582	3.5016	2.7627
		RBF	0.7719	1.9555	1.4909	0.7395	1.8417	1.4403
		BP	0.7834	1.9054	1.4729	0.7413	1.8351	1.4423
SNV	CARS	ELM	0.8823	2.7017	1.4989	0.7312	2.0423	1.4061
		LS-SVM	0.7044	2.2394	1.7338	0.6220	2.2871	1.7466
		RF	0.6455	0.2724	0.1936	0.0621	0.4268	0.3337
		RBF	0.7778	1.9300	1.5280	0.5086	2.5293	2.0673
		BP	0.7542	2.0298	1.4307	0.5517	2.4157	1.9926

Note: SG: savitzk–Golay. MSC: multiplicative scatter correction. SNV: standard normal variate. CARS: competitive adaptive reweighted sampling. ELM: extreme learning machine. LS-SVM: Least squares support vector machine. RF: random forest. RBF: radial basis function. BP: back propagation. R_C^2: the calibration set coefficient of determination. R_P^2: the prediction set coefficient of determination. RMSE: root mean square error. MAE: mean absolute error. The bold line represented the high prediction accuracy.

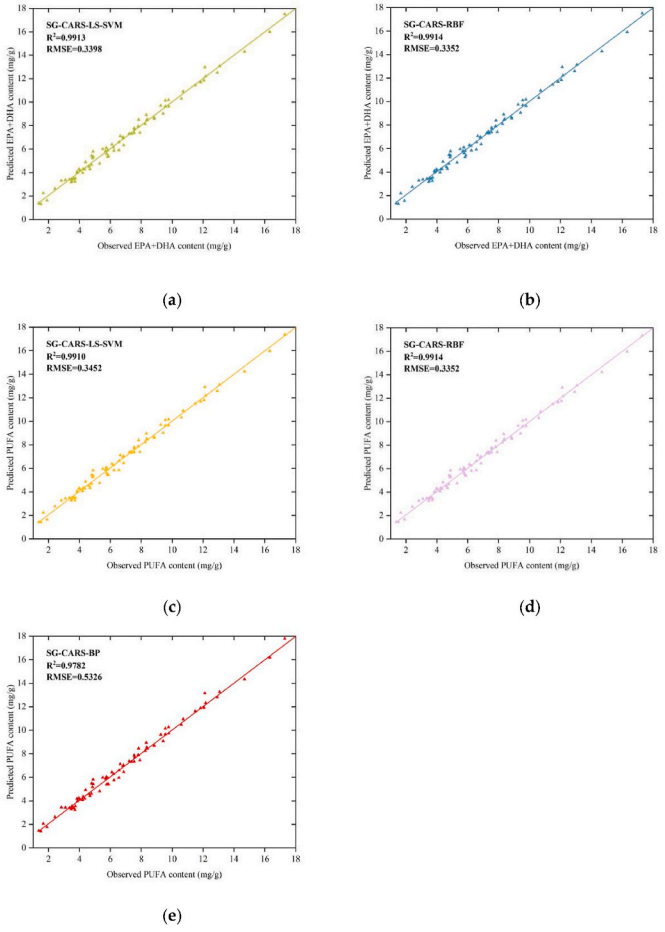


Fig. 5. Optimal prediction model for the contents of the EPA + DHA and PUFAs in common carp. (a) Muscle EPA + DHA prediction results using SG preprocessing method, CARS optimal wavelength screening method and LS-SVM model; (b) Muscle EPA + DHA prediction results using SG + CARS + RBF; (c) Muscle PUFA prediction results using SG + CARS + LS-SVM; (d) Muscle PUFA prediction results using SG + CARS + RBF; (e) Muscle PUFA prediction results using SG + CARS + BP. The numbers on the X-axis represent the observed values. The numbers on the Y-axis represent the predicted values.

method had very high consistency with the GC method and therefore offered a quick, non-destructive, and simple detection tool for measuring PUFAs in live fish fillets.

Although recent studies have reported the use of spectroscopic techniques and ML methods to predict muscle FAs contents, many disadvantages are still unsolved. Firstly, previous studies predominantly relied on fish muscular spectra to predict the muscular levels of FAs (Cheng et al., 2019; Ma et al., 2018; Wang et al., 2020; Lintvedt et al., 2023). Therefore, these methods also require the removal of the fish skin and are destructive. Secondly, most studies have concentrated on predicting the FAs content of meat using NIRS (700–2500 nm) and Raman spectroscopy (Xu et al., 2015; Ma and Sun, 2020; Wang et al., 2020; Afseth et al., 2022; Lintvedt et al., 2023). However, the detectable regions of the NIRS device were much smaller and the result could not be representative of the whole sample. Raman spectroscopy faces the challenge of being unsuitable for operation on the spot (Petersen et al., 2021; Lintvedt et al., 2023). In addition, most of the existing studies utilizing hyperspectral image data of the skin have focused on real-time monitoring of human skin characteristics for clinical assessment of diseases. However, these studies lacked quantitative analysis of the data (Rosen et al., 1990; Yudovsky et al., 2011; Ross et al., 2014; Paul et al., 2015). Zherebtsov et al. (2019) quantitatively assessed blood volume fractions and oxygen saturation by acquiring skin spectral information, but this process lacked the ability to provide an accurate fit coefficient for quantitative analysis. Thirdly, the prediction accuracy of muscle FAs contents could be improved, although the reported accuracy reached 0.9237 (Cheng et al., 2019; Ma et al., 2018; Wang et al., 2020). Our method overcomes these three disadvantages. To our knowledge, this is the first study to utilize skin spectral information to predict muscular PUFAs in fish. The HSI system could detect larger muscular regions than the NIRS device and is more suitable for detection on the spot than Raman spectroscopy. Finally, we obtained higher prediction accuracies for muscular PUFAs and EPA + DHA contents (both 0.9914) than did previous studies.

We demonstrated that the preprocessing method was an important factor affecting the prediction accuracy. The accuracies of the SG processed data were better than those of the other two methods. Pang et al. found that Raman spectroscopy combined with the SG method and employing the partial least squares (PLS) model yielded an R_P^2 value as high as 0.9836 (Pang et al., 2022). One possible reason for the better performance of the SG method is that it rateins useful spectral information with the noise eliminated (Fig. 3). However, the latter two

Table 4
Performance of modeling models to predict PUFAs contents using different preprocessing methods and characteristic wavelength screening methods.

Preprocessing Method	Band Selection	Modeling Method	Calibration			Prediction		
			R_C^2	RMSE	MAE	R_P^2	RMSE	MAE
SG	CARS	ELM	0.5880	2.6282	1.4887	0.4588	3.6383	2.7785
		LS-SVM	**0.9935**	**0.3312**	**0.2692**	**0.9910**	**0.3452**	**0.1192**
		RF	0.8270	1.7028	0.8401	0.6909	2.0060	1.3649
		RBF	**0.9934**	**0.3326**	**0.2704**	**0.9914**	**0.3346**	**0.2660**
		BP	**0.9934**	**0.3314**	**0.2651**	**0.9782**	**0.5326**	**0.4473**
MSC	CARS	ELM	0.6813	5.3420	1.2082	0.6604	4.8777	1.3804
		LS-SVM	0.9005	1.3111	0.9815	0.7387	1.8902	1.5006
		RF	0.5718	2.6792	1.5713	0.0329	3.5483	2.7738
		RBF	0.7675	1.9743	1.5055	0.7339	1.8614	1.4585
		BP	0.7653	1.9834	1.5308	0.7583	1.7739	1.3745
SNV	CARS	ELM	0.5880	2.6282	1.4887	0.4541	3.6383	2.7785
		LS-SVM	0.7498	2.0689	1.5974	0.6408	2.2369	1.7221
		RF	0.5880	2.6282	1.4887	0.5751	2.6689	1.4959
		RBF	0.7341	2.1115	1.6624	0.5708	2.3639	0.2365
		BP	0.6796	2.3177	1.6897	0.5108	2.5237	1.9471

Note: SG: savitzk–Golay. MSC: multiplicative scatter correction. SNV: standard normal variate. CARS: competitive adaptive reweighted sampling. ELM: extreme learning machine. LS-SVM: Least squares support vector machine. RF: random forest. RBF: radial basis function. BP: back propagation. R_C^2: the calibration set coefficient of determination. R_P^2: the prediction set coefficient of determination. RMSE: root mean square error. MAE: mean absolute error. The bold line represented the high prediction accuracy.

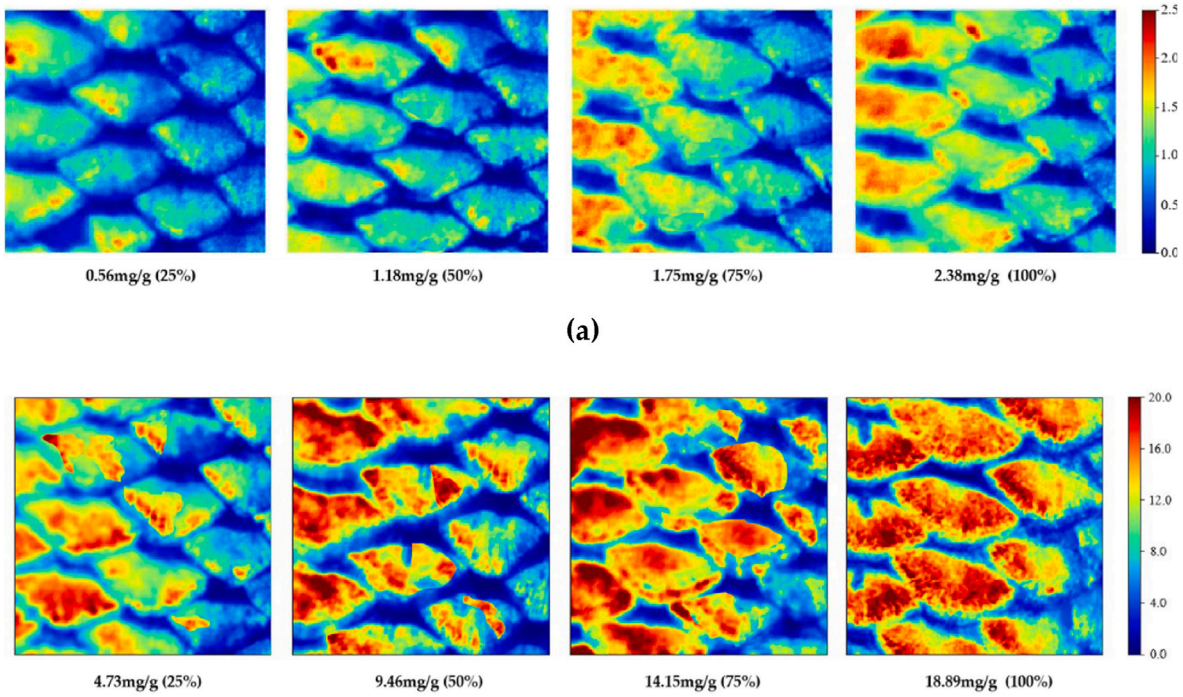


Fig. 6. Visualization of the distribution of EPA + DHA (a) and PUFAs (b) contents in common carp. The numbers on the right indicate the contents of the EPA + DHA and PUFAs and were scaled with different colors. The values under the samples indicate the EPA + DHA and PUFAs contents measured by the GC method. The color from blue to red represents the values from low content to high content. The percentages meant that the samples had the contents at the lowest quartile, median, upper quartile, and highest. (For interpretation of the references to color in this figure legend, the reader is referred to the Web version of this article.)

methods greatly compressed the spectral data ranging from 500 to 900 nm, resulting in the loss of spectral data.

The modeling method was the second factor affecting the prediction accuracy. The RBF model performed optimally in the prediction of both PUFAs and EPA + DHA contents ($R_P^2 = 0.9914$). The data is similar to the result of Yang et al. (2023) where they applied HSI (400–1000 nm) and the RBF prediction model to predict the FAs contents of camellia seed oil with an R_P^2 of 0.8437. Therefore, RBF modeling is possibly suitable for predicting nutrient contents.

In summary, our method presents several advantages. Firstly, it collected the fish skin spectral information, eliminating the requirement for muscle preparation. Hence, it is rapid, non-destructive, and inexpensive. Secondly, the equipment is portable for on-site assessment. Thirdly, it has another feature of high prediction accuracy. Finally, visual distribution maps are beneficial for intuitively assessing the nutritional quality of fish muscle. Overall, our method combining the skin

HSI and ML method is very promising for the massive, rapid, and non-destructive determination of PUFAs and EPA + DHA in fish fillets on the spot.

5. Conclusion

This study investigated the potential of utilizing the HSI technique and ML method for measuring muscular PUFAs contents with skin spectra. The combination of the SG preprocessing method and the CARS dimensionality reduction method effectively increased the prediction accuracy. The results demonstrated that the RBF model with SG preprocessing based on selected characteristic wavelengths achieved optimal predictive power for estimating the EPA + DHA ($R_P^2 = 0.9914$ and RMSE = 0.3352) and PUFAs contents ($R_P^2 = 0.9914$ and RMSE = 0.3346). Furthermore, the distributions of the EPA + DHA and PUFA contents provide more detailed and intuitive information about common carp fillet quality. Compared with the GC method and previous spectroscopic techniques, our method generated many improvements and is suitable for the massive, rapid, and non-destructive determination of fish fillet nutrients.

CRediT authorship contribution statement

Yi-Ming Cao: Methodology, Software, Validation, Formal analysis, Investigation, Data curation, Writing – original draft, Visualization. **Yan Zhang:** Methodology, Investigation, Project administration, Funding acquisition. **Qi Wang:** Investigation. **Ran Zhao:** Investigation. **Mingxi Hou:** Investigation. **Shuang-Ting Yu:** Investigation, Data curation. **Kai-Kuo Wang:** Investigation. **Ying-Jie Chen:** Investigation. **Xiao-Qing Sun:** Investigation. **Shijing Liu:** Validation. **Jiong-Tang Li:** Conceptualization, Methodology, Validation, Resources, Writing – review & editing, Supervision, Project administration, Funding acquisition.

Ethics statement

This study was approved by the recommendations of the Animal Care and Use Committee of the Chinese Academy of Fishery Sciences.

Declaration of competing interest

The authors declare that they have no conflicts of interest to declare for this publication.

Acknowledgments

This research was funded by the National Key Research and Development Program (grant number 2023YFD240203), the Special Scientific Research Funds for Central Non-profit Institutes, the Chinese Academy of Fishery Sciences (grant number 2023TD25), the Central Public-interest Scientific Institution Basal Research Fund, CAFS (grant number 2023A003), and the National Freshwater Genetic Resource Centre (grant number FGRC: 18537).

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.crfs.2024.100929.

Data availability

Data will be made available on request.

References

Abbas, O., Dardenne, P., Baeten, V., 2012. Near-infrared, mid-infrared, and Raman spectroscopy. In: Picó, Y. (Ed.), Chemical Analysis of Food: Techniques and Applications. Elsevier Inc, Waltham, MA, pp. 59–91. https://doi.org/10.1016/B978-0-12-384862-8.00003-0.

Afseth, N.K., Dankel, K., Andersen, P.V., Difford, G.F., Horn, S.S., Sonesson, A., Hillestad, B., Wold, J.P., Tengstrand, E., 2022. Raman and near infrared spectroscopy for quantification of fatty acids in muscle tissue-A salmon case study. Foods 11 (7), 962. https://doi.org/10.3390/foods11070962.

Aiyelokun, O.O., Agbede, O.A., 2021. Development of random forest model as decision support tool in water resources management of Ogun headwater catchments. Appl. Water Sci. 11, 1–9. https://doi.org/10.1007/s13201-021-01461-x.

Berhe, D.T., Eskildsen, C.E., Lametsch, R., Hviid, M.S., van den Berg, F., Engelsen, S.B., 2016. Prediction of total FAs parameters and individual fatty acids in pork backfat using Raman spectroscopy and chemometrics: understanding the cage of covariance between highly correlated fat parameters. Meat Sci. 111, 18–26. https://doi.org/10.1016/j.meatsci.2015.08.009.

Cao, Y.M., Zhang, Y., Yu, S.T., Wang, K.K., Chen, Y.J., Xu, Z.M., Ma, Z.Y., Chen, H.L., Wang, Q., Zhao, R., Sun, X.Q., Li, J.T., 2023. Rapid and non-invasive assessment of texture profile analysis of common carp (*Cyprinus carpio* L.) using hyperspectral imaging and machine learning. Foods 12 (17), 3154. https://doi.org/10.3390/foods12173154.

Chapman, J., Elbourne, A., Truong, V.K., Cozzolino, D., 2020. Shining light into meat - a review on the recent advances in *in invo* and carcass applications of near infrared spectroscopy. Int. J. Food Sci. Technol. 55 (3), 935–941. https://doi.org/10.1111/ijfs.14367.

Chen, C., Li, W., Su, H., Liu, K., 2014. Spectral-Spatial classification of hyperspectral image based on kernel extreme learning machine. Rem. Sens. 6 (6), 5795–5814. https://doi.org/10.3390/rs6065795.

Cheng, L., Liu, G., He, J., Wan, G., Ma, C., Ban, J., Ma, L., 2020. Non-destructive assessment of the myoglobin content of Tan sheep using hyperspectral imaging. Meat Sci. 167, 107988. https://doi.org/10.1016/j.meatsci.2019.107988.

Cheng, J.H., Sun, D.W., 2017. Partial least squares regression (PLSR) applied to NIR and HSI spectral data modeling to predict chemical properties of fish muscle. Food Eng. Rev. 9, 36–49. https://doi.org/10.1007/s12393-016-9147-1.

Cheng, J.H., Sun, D.W., Liu, G., Chen, Y.N., 2019. Developing a multispectral model for detection of docosahexaenoic acid (DHA) and eicosapentaenoic acid (EPA) changes in fish fillet using physarum network and genetic algorithm (PN-GA) method. Food Chem. 270, 181–188. https://doi.org/10.1016/j.foodchem.2018.07.013.

Cheng, J., Sun, J., Xu, M., Zhou, X., 2023. Nondestructive detection of lipid oxidation in frozen pork using hyperspectral imaging technology. J. Food Compos. Anal. 123, 105497. https://doi.org/10.1016/j.jfca.2023.105497.

Chiu, H.H., Kuo, C.H., 2020. Gas chromatography-mass spectrometry-based analytical strategies for fatty acid analysis in biological samples. J. Food Drug Anal. 28 (1), 60–73. https://doi.org/10.1016/j.jfda.2019.10.003.

De Carvalho, C.C., Caramujo, M.J., 2018. The various roles of fatty acids. Molecules 23 (10), 2583. https://doi.org/10.3390/molecules23102583.

Djuricic, I., Calder, P.C., 2021. Beneficial outcomes of omega-6 and omega-3 polyunsaturated fatty acids on human health: an update for 2021. Nutrients 13 (7), 2421. https://doi.org/10.3390/nu13072421.

Ecker, J., Scherer, M., Schmitz, G., Liebisch, G., 2012. A rapid GC–MS method for quantification of positional and geometric isomers of fatty acid methyl esters. J. Chromatogr. B 897, 98–104. https://doi.org/10.1016/j.jchromb.2012.04.015.

He, Q., Wang, R.K., 2019. Analysis of skin morphological features and real-time monitoring using snapshot hyperspectral imaging. Biomed. Opt Express 10 (11), 5625–5638. https://doi.org/10.1364/BOE.10.005625.

Jamshidi, A., Cao, H., Xiao, J., Simal-Gandara, J., 2020. Advantages of techniques to fortify food products with the benefits of fish oil. Food Res. Int. 137, 109353. https://doi.org/10.1016/j.foodres.2020.109353.

Kamruzzaman, M., Makino, Y., Oshita, S., 2016. Hyperspectral imaging for real-time monitoring of water holding capacity in red meat. LWT–Food Sci. Technol. 66, 685–691. https://doi.org/10.1016/j.lwt.2015.11.021.

Khan, A., Vibhute, A.D., Mali, S., Patil, C.H., 2022. A systematic review on hyperspectral imaging technology with a machine and deep learning methodology for agricultural applications. Ecol. Inf. 69, 101678. https://doi.org/10.1016/j.ecoinf.2022.101678.

Kamper, W., Trueman, S.J., Tahmasbian, I., Bai, S.H., 2020. Rapid determination of nutrient concentrations in hass avocado fruit by vis/nir hyperspectral imaging of flesh or skin. Rem. Sens. 12 (20), 3409. https://doi.org/10.3390/rs12203409.

Li, H., Liang, Y., Xu, Q., Cao, D., 2009. Key wavelengths screening using competitive adaptive reweighted sampling method for multivariate calibration. Anal. Chim. Acta 648 (1), 77–84. https://doi.org/10.1016/j.aca.2009.06.046.

Lintvedt, T.A., Andersen, P.V., Afseth, N.K., Heia, K., Lindberg, S.K., Wold, J.P., 2023. Raman spectroscopy and NIR hyperspectral imaging for in-line estimation of fatty acid features in salmon fillets. Talanta 254, 124113. https://doi.org/10.1016/j.talanta.2022.124113.

Ma, J., Sun, D.W., 2020. Prediction of monounsaturated and polyunsaturated fatty acids of various processed pork meats using improved hyperspectral imaging technique. Food Chem. 321, 126695. https://doi.org/10.1016/j.foodchem.2020.126695.

Ma, J., Pu, H., Sun, D.W., 2018. Predicting intramuscular fat content variations in boiled pork muscles by hyperspectral imaging using a novel spectral pre-processing technique. Lwt. 94, 119–128. https://doi.org/10.1016/j.lwt.2018.04.030.

Pang, L., Chen, H., Yin, L., Cheng, J., Jin, J., Zhao, H., Liu, Z., Dong, L., Yu, H., Lu, X., 2022. Rapid fatty acids detection of vegetable oils by Raman spectroscopy based on competitive adaptive reweighted sampling coupled with support vector regression. Food Qual. Saf. 6 (4), 10. https://doi.org/10.1093/fqsafe/fyac053.

Paul, D.W., Ghassemi, P., Ramella-Roman, J.C., Prindeze, N.J., Moffatt, L.T., Alkhalil, A., Shupp, J.W., 2015. Noninvasive imaging technologies for cutaneous wound assessment: a review. Wound Repair Regen. 23 (2), 149–162. https://doi.org/10.1111/wrr.12262.

Petersen, M., Yu, Z., Lu, X., 2021. Application of Raman spectroscopic methods in food safety: a review. Biosensors 11 (6), 187. https://doi.org/10.3390/bios11060187.
Ravikanth, L., Jayas, D.S., White, N.D., Fields, P.G., Sun, D.W., 2017. Extraction of spectral information from hyperspectral data and application of hyperspectral imaging for food and agricultural products. Food Bioprocess Technol. 10, 1–33. https://doi.org/10.1007/s11947-016-1817-8.
Rohman, A., Putri, A.R., Windarsih, A., Nisa, K., Lestari, L., 2021. The employment of analytical techniques and chemometrics for authentication of fish oils: a review. Food Control 124, 107864. https://doi.org/10.1016/j.foodcont.2021.107864.
Rosen, C.F., Jacques, S.L., Stuart, M.E., Gange, R.W., 1990. Immediate pigment darkening: visual and reflectance spectrophotometric analysis of action spectrum. Photochem. Photobiol. 51 (5), 583–588. https://doi.org/10.1111/j.1751-1097.1990.tb01969.x.
Ross, C.G., Chwal, C., Beckstead, J.A., Byard, R.W., Langlois, N.E., 2014. Hyperspectral imaging of bruises. Pathology 46, S88. https://doi.org/10.1097/01.PAT.0000443642.14560.55.
Sadeghin, B., Sarrafzadeh, M.H., Jin, J., Dupre, C., Watanabe, M., Legrand, J., Grizeau, D., 2018. Variation of fatty acids composition in the hydrocarbon producer Botryococcus braunii BOT 22. Biomass Bioenergy 119, 456–461. https://doi.org/10.1016/j.foodchem.2009.12.089.
Saha, D., Manickavasagan, A., 2021. Machine learning techniques for analysis of hyperspectral images to determine quality of food products: a review. Curr. Res. Food Sci. 4, 28–44. https://doi.org/10.1016/j.crfs.2021.01.002.
Sun, Y., Yuan, M., Liu, X., Su, M., Wang, L., Zeng, Y., Zang, H., Nie, L., 2021. A sample selection method specific to unknown test samples for calibration and validation sets based on spectra similarity. Spectrochim. Acta, Part A 258, 119870. https://doi.org/10.1016/j.saa.2021.119870.
Sushchik, N.N., Gladyshev, M.I., Kalachova, G.S., 2007. Seasonal dynamics of fatty acid content of a common food fish from Yenisei river, Siberian grayling, Thymallus arcticus. Food Chem. 104 (4), 1353–1358. https://doi.org/10.1016/j.foodchem.2007.01.050.
Tahmasbian, I., Wallace, H.M., Gama, T., Bai, S.H., 2021. An automated non-destructive prediction of peroxide value and free fatty acid level in mixed nut samples. Lebensm. Wiss. Technol. 143, 110893. https://doi.org/10.1016/j.lwt.2021.110893.
Wang, C., Wang, S., He, X., Wu, L., Li, Y., Guo, J., 2020. Combination of spectra and texture data of hyperspectral imaging for prediction and visualization of palmitic acid and oleic acid contents in lamb meat. Meat Sci. 169, 108194. https://doi.org/10.1016/j.meatsci.2020.108194.
Wu, D., Sun, D.W., 2013. Potential of time series-hyperspectral imaging (TS-HSI) for non-invasive determination of microbial spoilage of salmon flesh. Talanta 111, 39–46. https://doi.org/10.1016/j.talanta.2013.03.041.
Xia, L., Yang, J., Su, R., Zhou, W., Zhang, Y., Zhong, Y., Huang, S., Chen, Y., Li, G., 2019. Recent progress in fast sample preparation techniques. Anal. Chem. 92 (1), 34–48. https://doi.org/10.1021/acs.analchem.9b04735.
Xu, J.L., Riccioli, C., Sun, D.W., 2015. An overview on nondestructive spectroscopic techniques for lipid and lipid oxidation analysis in fish and fish products. Compr. Rev. Food Sci. Food Saf. 14 (4), 466–477. https://doi.org/10.1111/1541-4337.12138.
Yang, X., Jiang, P., Luo, Y., Shi, Y., 2023. Non-destructive detection of fatty acid content of camellia seed based on hyperspectral. J. Oleo Sci. 72 (1), 69–77. https://doi.org/10.5650/jos.ess22139.
Yu, H.D., Qing, L.W., Yan, D.T., Xia, G., Zhang, C., Yun, Y.H., Zhang, W., 2021. Hyperspectral imaging in combination with data fusion for rapid evaluation of tilapia fillet freshness. Food Chem. 348, 129129. https://doi.org/10.1016/j.foodchem.2021.129129.
Yudovsky, D., Nouvong, A., Schomacker, K., Pilon, L., 2011. Monitoring temporal development and healing of diabetic foot ulceration using hyperspectral imaging. J. Biophot. 4 (7–8), 565–576. https://doi.org/10.1002/jbio.201000117.
Zhang, G., Abdulla, W., 2022. New Zealand honey botanical origin classification with hyperspectral imaging. J. Food Compos. Anal. 109, 104511. https://doi.org/10.1016/j.jfca.2022.104511.
Zhang, T.T., Xu, J., Wang, Y.M., Xue, C.H., 2019. Health benefits of dietary marine DHA/EPA-enriched glycerophospholipids. Prog. Lipid Res. 75, 100997. https://doi.org/10.1016/j.plipres.2019.100997.
Zhang, Y., Sun, X.Q., Ye, Y.Q., Wang, Q., Li, Q.S., Zhao, R., Wang, H.W., Li, J.T., 2021. Association between the polymorphisms of fads2a and fads2b and poly-unsaturated fatty acids in common carp (*Cyprinus carpio*). Animals. 11 (6), 1780. https://doi.org/10.3390/ANI11061780.
Zheng, X., Li, Y., Wei, W., Peng, Y., 2019. Detection of adulteration with duck meat in minced lamb meat by using visible near-infrared hyperspectral imaging. Meat Sci. 149, 55–62. https://doi.org/10.1016/j.meatsci.2018.11.005.
Zherebtsov, E., Dremin, V., Popov, A., Doronin, A., Kurakina, D., Kirillin, M., Meglinski, I., Bykov, A., 2019. Hyperspectral imaging of human skin aided by artificial neural networks. Biomed. Opt Express 10 (7), 3545–3559. https://doi.org/10.1364/BOE.10.003545.