ARTICLE

BKCS BULLETIN OF THE
KOREAN CHEMICAL SOCIETY

Discrimination of omega-3 fatty acid oil forms by combining NMR spectroscopy with artificial intelligence

Neulhwi Yeo[1] ⓘ | **Jung Min Han**[1] ⓘ | **Mi Gang Kim**[1] ⓘ | **Jin Young Kim**[1] ⓘ |
Hyojin Cho[1] ⓘ | **Seon Yeong Lee**[1] ⓘ | **Joong-Hyuck Auh**[2] ⓘ | **Byung Hee Kim**[3] ⓘ |
Sangdoo Ahn[1] ⓘ

[1]Department of Chemistry, Chung-Ang University, Seoul, Republic of Korea

[2]Department of Food Science and Technology, Chung-Ang University, Anseong, Republic of Korea

[3]Department of Food and Nutrition, Sookmyung Women's University, Seoul, Republic of Korea

Correspondence
Sangdoo Ahn, Department of Chemistry, Chung-Ang University, Seoul 06974, Republic of Korea.
Email: sangdoo@cau.ac.kr

Abstract
This study presents an approach for discriminating omega-3 fatty acid forms using proton nuclear magnetic resonance (^{1}H-NMR) spectroscopy combined with machine learning and deep learning techniques. A total of 90 samples, comprising triglyceride, re-esterified triglyceride, and ethyl ester forms, were analyzed. Principal component analysis–linear discriminant analysis, support vector machine (SVM), artificial neural network (ANN), and one-dimensional convolutional neural network (1D CNN) models were applied using binned spectral data. In contrast, a two-dimensional convolutional neural network (2D CNN) was constructed using spectral images. To prevent overfitting and optimize model hyperparameters, early stopping, cross-validation, and Bayesian optimization were used across the different machine learning and deep learning models. The 1D and 2D CNN models both achieved 100% accuracy on the training and test sets, while the SVM and ANN models yielded slightly lower but still excellent performance, with a test accuracy of 94.4%. Model interpretability was enhanced through SHapley Additive exPlanations and Gradient-weighted Class Activation Mapping, which identified critical spectral regions associated with classification decisions. These results demonstrate that the integration of artificial intelligence techniques with ^{1}H-NMR spectroscopy enables accurate, interpretable discrimination of omega-3 fatty acid forms, offering a promising strategy for supplement authentication and quality control.

KEYWORDS
artificial intelligence, deep learning-based CNN, fatty acids discrimination, NMR, re-esterified omega-3

INTRODUCTION

Omega-3 oils are edible oils rich in omega-3 fatty acids, notably eicosapentaenoic acid (EPA) and docosahexaenoic acid (DHA), which have been widely used in dietary supplements owing to their numerous health benefits.[1] Extensive research has demonstrated that omega-3 fatty acids contribute to supporting cardiovascular, brain, and ocular health, as well as mitigating inflammatory responses.[2–4] Reflecting growing consumer awareness of these benefits, the global omega-3 oil supplements market size reached approximately USD 7.3 billion in 2022.[5]

Commercially, omega-3 supplements are available in three major chemical forms: triglycerides (TG), ethyl esters

(EE), and re-esterified triglycerides (rTG), each differing in molecular structure, bioavailability, and manufacturing process. Among these, rTG forms have garnered increasing attention due to their superior absorption efficiency and higher EPA and DHA contents compared to conventional TG forms.[6] The enhanced bioavailability and elevated concentrations of beneficial fatty acids make rTG products highly desirable in the marketplace. rTG oils are typically produced using a two-step chemical process involving the transesterification of TG to EE, followed by re-esterification with glycerol. However, this multi-step synthesis increases production costs, resulting in higher retail prices for rTG products compared to TG or EE forms. Furthermore, the complexity of the process often leads to incomplete conversion, leaving residual monoacylglycerols, diacylglycerols, and unreacted EE in the final product, thereby complicating chemical composition.[6–8] The high value of rTG products, combined with the absence of standardized regulatory definitions, raises the risk of economically motivated adulteration, wherein TG or EE forms may be artificially introduced to mimic or substitute authentic rTG oils. These factors underscore the urgent need for reliable analytical methods to authenticate and ensure the quality of omega-3 supplements.

Various analytical techniques have been developed to assess the composition and authenticity of omega-3 oils. Gas chromatography with flame ionization detection (GC-FID) remains the standard method for quantifying fatty acid profiles and ensuring regulatory compliance.[5,9,10] However, GC-FID typically requires derivatization, involves lengthy sample preparation and analysis times, and often lacks the structural specificity necessary to distinguish between molecular forms. In contrast, proton nuclear magnetic resonance (^{1}H-NMR) spectroscopy provides a rapid, non-destructive alternative with minimal sample preparation and has been extensively applied to both qualitative and quantitative analyses of omega-3 oils.[11,12] Additionally, ^{13}C-NMR, ^{31}P-NMR, and two-dimensional (2D) NMR techniques have been used for more detailed compositional profiling.[13,14]

Multivariate statistical methods, such as partial least squares discriminant analysis (PLS-DA) and orthogonal projections to latent structures discriminant analysis (OPLS-DA), have been introduced to enhance classification based on NMR data. For example, Amorim et al. successfully distinguished TG and EE forms using ^{1}H-NMR spectra combined with supervised learning techniques.[15] However, due to their hybrid composition, these models have not yet been adapted to identify rTG forms, which are structurally more complex than TG and EE forms.

As the complexity of food matrices and chemical compositions has increased, the need for more powerful and flexible analytical approaches has also increased. Consequently, recent studies have increasingly explored the use of artificial intelligence (AI) techniques for interpreting spectroscopic data in food analysis. Machine learning (ML) models have been successfully applied to a wide range of spectral data, including Raman, infrared (IR), and NMR.[16–18] In the context of omega-3 oils, for instance, support vector machine (SVM) algorithms have been used to classify certified omega-3 oils using Raman spectroscopy.[19]

This study builds upon previous efforts by exploring whether AI—particularly deep learning (DL)—can be used to identify the structural fingerprints of TG, EE, and rTG omega-3 fatty acid oils directly from ^{1}H-NMR spectra. Using a dataset comprising both commercial and experimental samples, this study applies multiple AI models—including SVM, principal component analysis-linear discriminant analysis (PCA–LDA), artificial neural networks (ANNs), and convolutional neural networks (CNNs)—to classify omega-3 oil forms. To ensure model interpretability, SHapley Additive exPlanations (SHAP) and Gradient-weighted Class Activation Mapping (Grad-CAM) are employed to visualize the key spectral regions influencing the classification decision. Through this integrated approach, the study not only advances methods for the detection and authentication of omega-3 supplements but also contributes to the development of interpretable, data-driven tools for quality assessment in food analysis.

EXPERIMENTS

Materials

A total of 90 commercial omega-3 supplement samples were analyzed in this study, comprising 30 samples each of rTG, TG, and EE products. The TG and EE standards (EE-DHA and EE-EPA) were purchased from Sigma-Aldrich (Seoul, Korea). In contrast, due to the limited availability of commercially standardized rTG materials, raw rTG oil was obtained in collaboration with a domestic manufacturing company (KD Pharma Co.) and used as a reference standard. For rTG supplement samples, only products that met the labeled omega-3 content and contained ≥70% omega-3 fatty acids (EPA and DHA) relative to total fatty acids were included in this classification analysis. Although rTG is theoretically composed entirely of omega-3 fatty acids, residual TG or EE forms are typically present due to limitations in the production process. Without a standardized definition for rTG form, a 70% threshold was adopted as the operational criterion. The reference rTG material used as a standard contained approximately 80% omega-3 fatty acids (EPA and DHA) based on total fatty acid composition.

Deuterated chloroform (CDCl$_3$, ≥99.8 atom% D) was obtained from BK Instruments (Daejeon, Korea) and used as the solvent for direct ^{1}H-NMR analysis of omega-3 samples without further purification. For NMR measurements, 20 μL of the omega-3 oil extracted from each capsule was dissolved in 600 μL of CDCl$_3$, and the resulting solution was transferred into a 5 mm (outer diameter) NMR tube for analysis.

^{1}H-NMR experiment and data processing

All NMR measurements were performed using a Varian 600 MHz NMR spectrometer (Palo Alto, CA, USA) equipped with a 5 mm PFG dual broadband 600NB probe and controlled by VnmrJ software (version 3.2A). Spectra were acquired with 64 k data points and a spectral width of 9615.4 Hz. The acquisition parameters were as follows: temperature, 25°C; pulse angle, 45°; acquisition time, 2 s; number of scans, 16; relaxation delay, 10 s. Chemical shifts were referenced to the residual proton signal of $CDCl_3$ at 7.26 ppm. Spectral preprocessing, including phase and baseline correction, was performed using Mnova software (Mestrelab Research, Santiago de Compostela, Spain).

For all classification models, the spectral region from 0.5 to 6.0 ppm—excluding the solvent signal—was selected as the analytical window. For PCA-LDA, SVM, ANN, and 1D CNN models, the ^{1}H-NMR data were binned using a bin size of 0.005 ppm to reduce dimensionality and enhance model interpretability. In contrast, for 2D CNN models, the spectral data within the selected range were converted into grayscale images and saved in PNG format. Image preprocessing included black-and-white inversion, resizing, and interpolating using the OpenCV library (cv2.INTER_AREA). All spectral images were standardized to a resolution of 1024 × 128 pixels to ensure consistent input dimensions for model training.

A total of 90 ^{1}H-NMR spectra, consisting of 30 samples each for TG, EE, and rTG forms, were used for model development. The dataset was randomly divided into training and test sets using an 8:2 ratio, resulting in 72 samples for training and 18 samples for testing. This data split was consistently applied across all models to ensure comparability of classification performance.

AI classification models

To discriminate the chemical forms of omega-3 fatty acids, a series of ML and DL models were developed based on ^{1}H-NMR spectral data. All models were implemented using the following Python packages (v3.11.12) in the Google Colab environment: scikit-learn (v1.4.2), TensorFlow (v2.15.0), and Keras (v2.15.0). For 2D-CNN models, spectral images were assessed using OpenCV (v4.10.0). SHAP and Grad-CAM techniques were applied to interpret model decisions. This provided visual and quantitative insights into the spectral regions critical for classification. SHAP values were calculated using the SHAP library (v0.41.0), while Grad-CAM was implemented using TensorFlow v2.15.0.

PCA-LDA was employed as a classical multivariate method combining unsupervised and supervised learning.[20] PCA reduced spectral dimensionality while preserving variance, and LDA was used for class prediction based on the transformed features. PCA enabled clearer pattern recognition in the NMR spectra by projecting them onto orthogonal axes, whereas LDA maximized class separation by optimizing inter-class and intra-class variances.[21] The number of principal components used in LDA was optimized using five-fold cross-validation. SHAP was used to assess which spectral regions (ppm values) most contributed to classification performance.[22]

SVM models were implemented as a supervised classification approach that seeks an optimal hyperplane, maximizing the margin between different classes.[23] Several kernel functions were tested, including radial basis function (RBF), polynomial, and sigmoid. To optimize model performance, a randomized grid search approach was employed in combination with five-fold cross-validation to tune hyperparameters such as the kernel type, cost (C), and gamma (γ). C controlled the balance between margin maximization and classification error minimization, while γ defined the influence of individual training samples. Proper tuning of these parameters was essential to balance model generalizability and overfitting. SHAP analysis was also used to support feature attribution.

ANN models were implemented as fully connected feedforward networks capable of modeling complex non-linear relationships in the spectral data.[24] The input consisted of binned spectral vectors. Furthermore, the network architecture comprised an input layer, two hidden layers with ReLU activation functions, and an output layer with softmax activation for three-class classification. Bayesian optimization was used to tune hyperparameters such as the number of dense units and learning rate. SHAP analysis was conducted to identify the spectral regions that contributed most to class differentiation, providing interpretability comparable to other models.

CNNs were implemented to capture distinctive patterns in the spectral data. The 1D CNN processed binned spectral vectors, enabling the detection of local signal features along the chemical shift axis. In contrast, the 2D CNN interpreted the spectra as 2D images, allowing for the extraction of spatial features through 2D convolutional filters. Both architectures consisted of stacked convolutional and pooling layers, followed by fully connected layers, facilitating hierarchical feature extraction.[25] Hyperparameter optimization was performed using Bayesian optimization, targeting parameters such as the number of filters, convolutional layers, neurons in the fully connected layers, and the learning rate.[26] Early stopping was applied to prevent overfitting during training. Although the CNN models required higher computational resources compared to conventional ML models, they effectively captured structural and distributional patterns in the spectral data. To enhance model interpretability, Grad-CAM was employed to visualize the specific regions of the spectral images that contributed most significantly to the classification, complementing the numerical feature importance ranking provided by SHAP.[27]

Figure 1 provides a schematic summary of the overall analytical workflow, incorporating the procedures
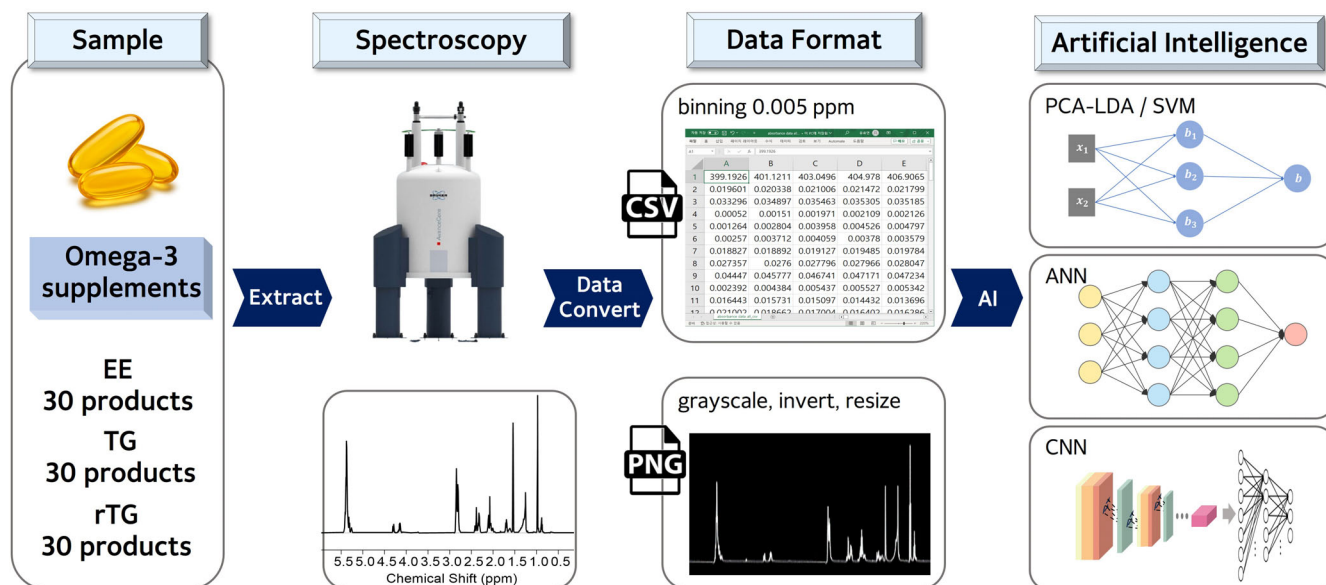
FIGURE 1 Schematic illustration of the overall process employed in this study.

described above, from sample preparation through spectral analysis and AI-based classification.

RESULTS AND DISCUSSION

^{1}H-NMR spectra of different omega-3 forms

A representative ^{1}H-NMR spectrum used for discriminating the forms of omega-3 samples extracted with CDCl$_3$ is shown in Figure 2, and the corresponding peak assignments are summarized in Table 1.[15] The chemical shift range of interest spans from 0.5 to 6.0 ppm, within which all relevant signals from omega-3 glycerides and fatty acids are observed.

Characteristic spectral differences among the three major forms of omega-3—TG, EE, and rTG—enable their differentiation by ^{1}H-NMR analysis. A key distinguishing feature lies in the methyl signals of saturated fatty acids. TG samples exhibit a relatively intense signal for terminal protons in the region of 0.81–0.93 ppm (peak 1), corresponding to saturated fatty acid methyl groups. Although these signals are theoretically absent in EE and rTG, trace amounts of saturated fatty acids—residual from incomplete purification—can produce subtle signals in the same region in rTG samples. In contrast, terminal protons of omega-3 fatty acids in EE and rTG appear downfield at 0.93–1.01 ppm (peak 2), reflecting the characteristic shift of unsaturated chains.

Another differentiating factor is the structural configuration, whether glycerol or EE. A distinctive signal of the ethyl group (CH$_3$–CH$_2$–O–) emerges solely in EE in the range of 4.09–4.15 ppm (peak 12). Contrastingly, the sn-1,3 proton signals of glycerol are evident at 4.06–4.33 ppm in TG and rTG (peak 11). The two prevalent
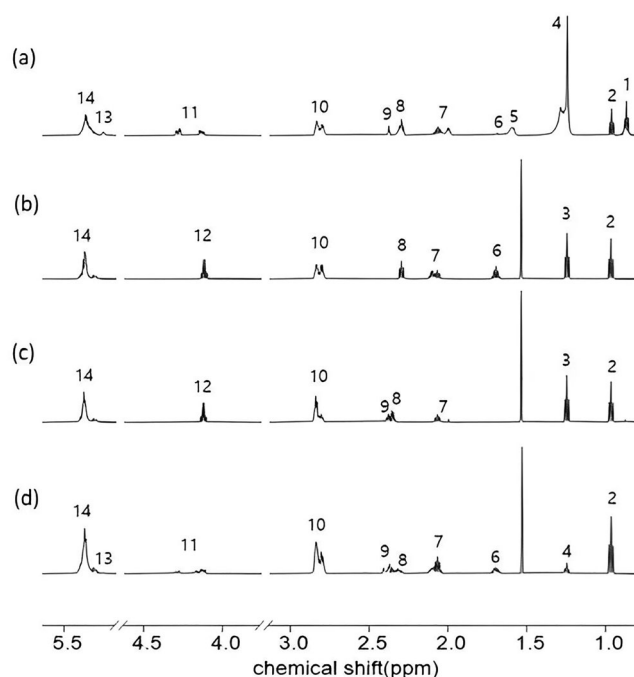

FIGURE 2 Representative ^{1}H-NMR spectra of different omega-3 forms: (a) TG, (b) EE-EPA, (c) EE-DHA, and (d) rTG.

forms of omega-3 fatty acids, EPA and DHA, also manifest distinct characteristic signals in the spectrum. The α-hydrogen signal of EPA appears in the range of 1.67–1.73 ppm (peak 6), while the α- and β-hydrogen signals are observed at 2.36–2.42 ppm (peak 9).

In summary, TG is characterized by peaks 1 and 11, EE exhibits peak 12, and rTG displays peak 11 in the ^{1}H-NMR spectrum. Peak 1 is either absent or of very low intensity in the spectra of EE and rTG, whereas peak 14, attributed

TABLE 1 Chemical shift assignment of the ^{1}H-NMR signals in CDCl$_3$ of the main protons of glycerides and fatty acids present in omega-3 supplements.

Peak no.	Chemical shift range (ppm)	Functional group	
		Type of proton	Compound
1	0.81–0.93	–CH$_3$	Saturated fatty acid
2	0.93–1.01	–CH$_3$	Omega-3 fatty acid
3	1.22–1.28	CH$_3$–CH$_2$–O–	EE
4	1.19–1.44	–CH$_2$–CH$_2$–CH$_2$–	Saturated fatty acid
5	1.55–1.66	–OCO–CH$_2$–CH$_2$–	Saturated fatty acid
6	1.67–1.73	–OCO–CH$_2$–CH$_2$–	EPA
7	2.03–2.13	CH$_2$–CH$_3$	Omega-3 fatty acid
8	2.26–2.26	–COC–CH$_2$–CH$_2$–	Fatty acid
9	2.36–2.42	–COC–CH$_2$–CH$_2$–	DHA
10	2.74–2.88	=CH–CH$_2$–CH=	Unsaturated fatty acid
11	4.06–4.33	ROCH$_2$–CH(OR′)–CH$_2$OR″	TG, rTG
12	4.09–4.15	CH$_3$–CH$_2$–O–	EE
13	5.22–5.28	ROCH$_2$–CH(OR′)–CH$_2$OR″	TG, rTG
14	5.28–5.45	–CH=CH–	Unsaturated fatty acid

to olefinic protons of unsaturated fatty acids, is typically more pronounced in these forms. These distinctive features of each omega-3 form are consistently observed across all the samples.

Discrimination of omega-3 fatty acid forms using AI models

All AI models—PCA-LDA, SVM, ANN, 1D CNN, and 2D CNN—were applied to the ^{1}H-NMR spectra of omega-3 supplements to classify different chemical forms: TG, rTG, and EE. The classification performance of each model is summarized in Table 2.

The PCA-LDA model achieved an accuracy of 97.2% for the training set (with two misclassifications among 72 samples) and 88.9% for the test set (with two misclassifications among 18 samples). While PCA-LDA effectively reduced dimensionality and captured the main variance in the spectral data, its linear nature limited its ability to fully separate the complex patterns among omega-3 forms, resulting in moderate performance compared to more sophisticated models. Optimized via randomized grid search, the SVM model yielded a training accuracy of 100% and a test accuracy of 94.4% (one misclassification in the test set). The model's ability to construct non-linear decision boundaries through kernel functions enhanced its discriminatory power, particularly in capturing subtle variations between structurally similar forms. The ANN model achieved a training accuracy of 100% and a test accuracy of 94.4%, with one misclassification in the test set. As a fully connected network, the ANN effectively captured nonlinear relationships within the spectral data. Although it lacks the spatial filtering capabilities of CNNs,

TABLE 2 Classification accuracy (%) of training and testing datasets for different AI models.

Data set	Accuracy (%)				
	PCA-LDA	SVM	ANN	1D CNN	2D CNN
Training	97.2	100	100	100	100
Test	88.9	94.4	94.4	100	100

its performance was comparable to that of the SVM model, underscoring the potential of neural networks to learn global spectral patterns even from limited datasets.

The 1D and 2D CNN models achieved 100% accuracy on both the training and test sets. The 2D CNN was trained using preprocessed spectral images, while the 1D CNN directly processed one-dimensional spectral vectors. Techniques such as batch normalization, early stopping, and Bayesian hyperparameter optimization were implemented to enhance model generalization and prevent overfitting. Both CNN architectures effectively captured local patterns within the spectral data—spatial features in 2D images and sequential patterns in 1D vectors—allowing for highly accurate discrimination among omega-3 forms.

SHAP and Grad-CAM analyses were conducted to further elucidate the basis for model decisions. In the PCA-LDA model, SHAP analysis identified the chemical shift region of 1.25–1.30 ppm (peak 4), corresponding to methylene protons of saturated fatty acids, as the most influential variable (Figure 3a). The presence of this feature, particularly in TG samples, contributed significantly to the classification. In the SVM model, the SHAP analysis similarly highlighted peak 4 but also identified peak 2 (0.975 ppm), corresponding to the terminal proton of
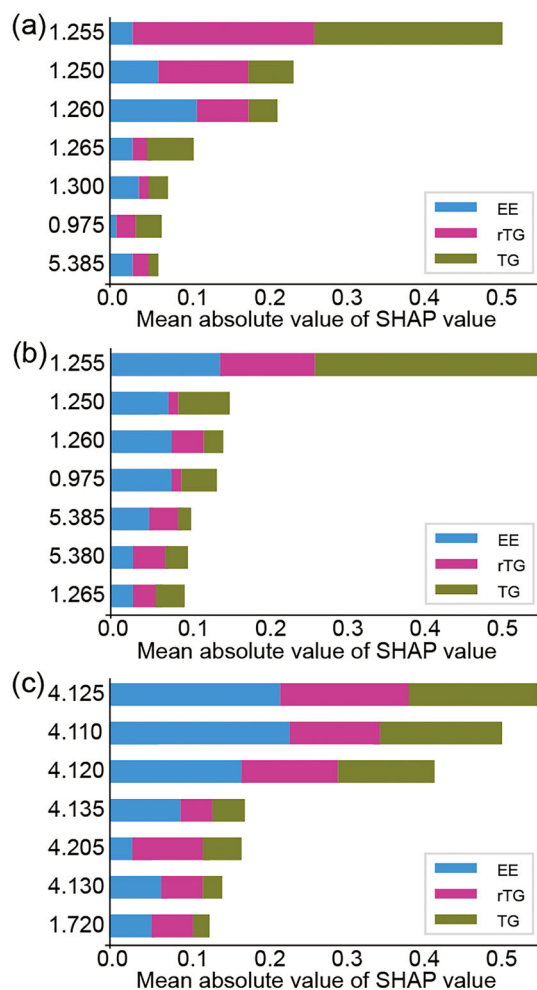
FIGURE 3 SHapley Additive exPlanations (SHAP) values indicating feature importance for oil classification using NMR spectra with (a) PCA-LDA, (b) SVM, and (c) ANN models. Only the top seven most influential features are displayed, with longer bars representing greater contributions to the model's classification decisions.

FIGURE 4 Grad-CAM results for (a) 1D and (b) 2D CNN models for each omega-3 form. Red-colored regions indicate the most influential spectral areas that contributed to the model's classification decision.

omega-3 fatty acids, and peak 14 (5.380–5.385 ppm), associated with olefinic protons of unsaturated fatty acids, as key variables (Figure 3b). This broader distribution of important features suggests that SVM captured more complex structural differences among the forms. In the ANN model, SHAP analysis highlighted the region between 4.11 and 4.20 ppm, which includes peaks 11 and 12. Peak 11 corresponds to glycerol backbone protons, characteristic of TG and rTG forms, while peak 12 corresponds to OCH_2 protons from the ethyl ester group, specific to the EE form. These findings indicate that the ANN model relied on subtle but structurally meaningful differences in this overlapping region to distinguish among the three forms (Figure 3c).

A comprehensive interpretation of the spectral features captured by the CNN models was obtained through Grad-CAM analysis (Figure 4), which generates class activation maps by leveraging gradient information to highlight the regions o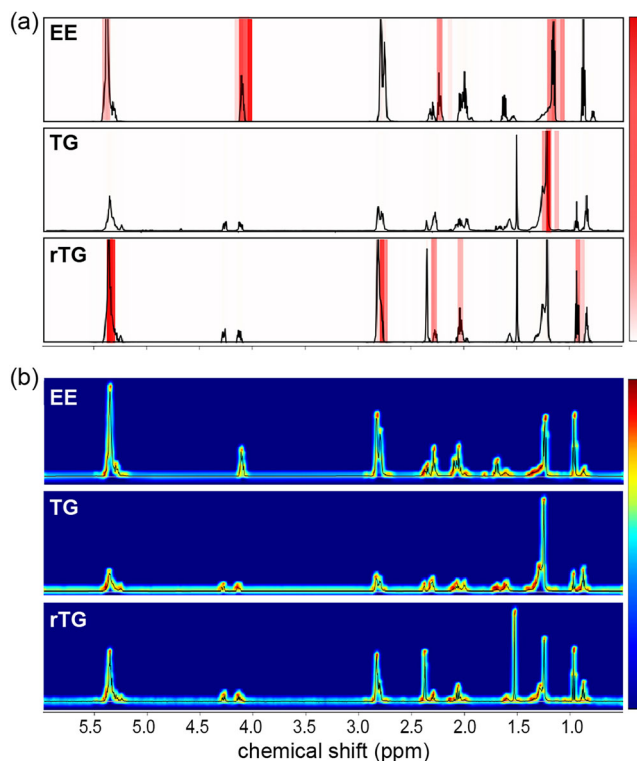f the spectral image that the model focuses on during classification. Grad-CAM analysis of the 1D CNN revealed distinct activation patterns in chemically meaningful spectral regions. For EE samples, strong activation was observed at 1.22–1.28 ppm (peak 3) and 4.09–4.10 ppm (peak 12), corresponding to ethyl ester protons that differentiate EE from the other forms. TG samples showed activation primarily in the saturated fatty acid methylene region (1.19–1.44 ppm), while rTG samples exhibited activation at 0.95–1.01 ppm, 2.74–2.88 ppm, and 5.28–5.45 ppm, reflecting features associated with unsaturated fatty acids and re-esterified structures.

The 2D CNN model exhibited broader activation patterns that included these critical regions as well as additional class-specific signals. In TG samples, prominent activation was noted in the methyl and methylene regions (0.81–1.44 ppm), along with other upfield aliphatic regions. In the case of rTG samples, the activation was primarily concentrated in the glycerol backbone region (4.06–4.33 ppm), corresponding to the *sn*-1,3 protons, while additional attention was also distributed across nearby regions related to the re-esterified configuration. For EE samples, distinct activation was detected in the terminal methyl region (0.93–1.01 ppm), the ethyl group region (4.09–4.15 ppm), and the olefinic region (5.28–5.45 ppm), indicating that the model identified multiple chemically relevant signals unique to this form. These findings, as revealed by Grad-CAM, demonstrate that the CNN model did not rely on a limited set of localized spectral

signals for classification, but instead captured class-specific spectral signatures by integrating information from multiple characteristic regions. This integrative pattern recognition likely contributed to the model's high classification accuracy across all omega-3 forms.

Comparison with previous studies further highlights the strengths of the current approach. Amorim et al. demonstrated omega-3 classification using ¹H-NMR and multivariate analysis; however, their study focused on only two forms (TG and EE), utilized a limited spectral region, and did not provide interpretability regarding key signals.[15] Kim et al. applied multivariate analyses based on lipid profiles obtained by liquid chromatography with evaporative light scattering detection (LC-ELSD) to authenticate rTG-type omega-3 oils and detect adulteration with EE-type oils.[5] However, their approach was based on targeted compositional analysis rather than direct spectroscopic fingerprinting and primarily aimed at detecting adulteration rather than comprehensive classification. In contrast, the current study classified all three major omega-3 forms using full-range ¹H-NMR spectra and explainable AI techniques, enabling both accurate and interpretable discrimination.

Despite the excellent classification performance of the CNN models, the relatively small sample size ($N = 90$) remains a key limitation. The classification task involved three chemically distinct and well-separated omega-3 forms (TG, EE, rTG), with ¹H-NMR spectra exhibiting high signal-to-noise ratios and informative structural features. These factors likely facilitated robust model learning despite the limited dataset. However, more complex tasks—such as those involving mixtures, degradation products, or overlapping spectral features—would require a larger and more diverse dataset to ensure model robustness and generalizability.

When working with small datasets, the application of deep learning models such as CNNs necessitates careful validation design to ensure generalizable performance. In this study, a practical and resource-efficient strategy was adopted, combining a shallow CNN architecture, Bayesian hyperparameter optimization with internal validation, early stopping, and evaluation using a fully independent test set (see Tables S3 and S4). This approach enabled effective model complexity control, guided tuning, and reliable performance assessment without the computational burden of repeated training cycles. Validation strategies should be selected based on dataset size, model architecture, and available computational resources to ensure credible results.

Model interpretability is also an essential consideration. Grad-CAM was used to visualize class-specific activation patterns in the CNN, highlighting spectral regions of chemical relevance. Although Grad-CAM provides coarser interpretive resolution compared to feature-level methods like SHAP, it confirmed that the CNN predictions were grounded in chemically meaningful features. For the present task, this level of interpretability may be sufficient. However, in more complex or regulated settings, additional or more advanced interpretability tools may be required to ensure transparency and confidence in model decisions.

Ultimately, model selection for spectroscopic analysis should not be based solely on classification accuracy. Other factors—such as data characteristics, model interpretability, computational efficiency, and optimization strategies—must also be considered. The approach presented here was designed to balance these elements within practical constraints, offering a framework that is both analytically robust and operationally feasible.

In summary, the integration of AI models with ¹H-NMR spectroscopy demonstrates considerable potential for the rapid and accurate classification of omega-3 fatty acid forms. Particularly, the application of CNN models to spectral data provides a powerful tool for distinguishing chemically similar species, offering new possibilities for food authenticity verification and quality control in the supplement industry.

CONCLUSION

In this study, AI techniques were applied to classify omega-3 supplement forms based on ¹H-NMR spectral data. ML and DL models, including PCA-LDA, SVM, ANN, 1D CNN, and 2D CNN, were developed and evaluated. Among the models tested, the CNN models demonstrated the highest performance, achieving 100% accuracy for both training and test sets. While the PCA-LDA model yielded comparatively lower accuracy, the SVM and ANN models also showed excellent performance, highlighting the effectiveness of both kernel-based and neural network-based approaches in spectral classification. SHAP and Grad-CAM analyses were employed to enhance interpretability. These explainability tools provided insight into the spectral regions most relevant to classification decisions, with Grad-CAM particularly highlighting the ability of the CNN models to localize chemically meaningful features across the full ¹H-NMR spectrum.

The integration of AI-based models with ¹H-NMR spectroscopy offers a rapid, accurate, and interpretable alternative to traditional spectral interpretation for omega-3 classification. This approach holds considerable promise for improving quality assurance, detecting adulteration, and verifying the authenticity of omega-3 supplements and related food products.

DATA AVAILABILITY STATEMENT
The data that support the findings of this study are available from the corresponding author upon reasonable request.

ORCID

Neulhwi Yeo https://orcid.org/0009-0007-4198-0032
Jung Min Han https://orcid.org/0009-0000-9750-9054
Mi Gang Kim https://orcid.org/0009-0009-6660-2913
Jin Young Kim https://orcid.org/0009-0005-0551-7214
Hyojin Cho https://orcid.org/0009-0004-4831-8349
Seon Yeong Lee https://orcid.org/0009-0004-0932-9731
Joong-Hyuck Auh https://orcid.org/0000-0001-9425-2599
Byung Hee Kim https://orcid.org/0000-0002-4599-6775
Sangdoo Ahn https://orcid.org/0000-0003-3803-9210

REFERENCES

[1] D. Swanson, R. Block, S. A. Mousa, *Adv. Nutr.* **2012**, *3*, 1.
[2] R. Crupi, S. Cuzzocrea, *Biomolecules* **2022**, *12*, 242.
[3] J. Li, B. L. R. Pora, K. Dong, J. Hasjim, *Food Sci. Nutr.* **2021**, *9*, 5229.
[4] D. Karageorgou, U. Rova, P. Christakopoulos, P. Katapodis, L. Matsakas, A. Patel, *Trends Food Sci. Technol.* **2023**, *136*, 169.
[5] H. Kim, E. Choi, J.-H. Auh, S. Ahn, H. S. Chun, B. H. Kim, *Food Chem.* **2025**, *463*, 141171.
[6] J. Dyerberg, P. Madsen, J. M. Møller, I. Aardestrup, E. B. Schmidt, *Prostaglandins, Leukotrienes Essent. Fatty Acids* **2010**, *83*, 137.
[7] G. F. Ferreira, J. G. B. Pessoa, L. F. R. Ríos Pinto, R. Maciel Filho, L. V. Fregolente, *Trends Food Sci. Technol* **2021**, *118*, 589.
[8] P. Lembke, in *Omega-6/3 Fatty Acids: Functions, Sustainability Strategies and Perspectives* (Eds: F. de Meester, R. R. Watson, S. Zibadi), Humana Press, Totowa, NJ **2013**, p. 353.
[9] S. M. Alinafiah, A. Azlan, A. Ismail, N.-K. M. A. Rashid, *Molecules* **2021**, *26*, 6592.
[10] C. T. Srigley, J. I. Rader, *J. Agric. Food Chem.* **2014**, *62*, 7268.
[11] J. Lv, C. Wang, X. Z. Zhang, Z. Lv, M. Yu, *J. Ocean Univ. China* **2020**, *19*, 1193.
[12] R. Sacchi, M. Savarese, L. Falcigno, I. Giudicianni, L. Paolillo, in *Modern Magnetic Resonance* (Ed: G. A. Webb), Springer, Dordrecht **2006**, p. 919.
[13] P. Dais, M. Misiak, E. Hatzakis, *Anal. Methods* **2015**, *7*, 5226.
[14] C. Remy, S. Danoun, M. Delample, C. Morris, V. Gilard, S. Balayssac, *Mag. Reson. Chem.* **2024**, *62*, 328.
[15] T. L. Amorim, Á. S. Granato, T. de Oliveira Mendes, M. A. L. de Oliveira, G. W. Amarante, M. A. de la Fuente, P. Gómez-Cortés, *J. Food Compost. Anal.* **2021**, *102*, 104060.
[16] H. Lim, S. Y. Lee, J. Y. Kim, Y. J. Shin, Y. Jang, H. Kim, B. H. Kim, S. Ahn, *Bull. Korean Chem. Soc.* **2025**, *46*, 131.
[17] L. Pan, P. Zhang, C. Daengngam, S. Peng, M. Chongcheawchamnan, *J. Raman Spectrosc.* **2022**, *53*, 6.
[18] B. H. Yun, H.-Y. Yu, H. Kim, S. Myoung, N. Yeo, J. Choi, H. S. Chun, H. Kim, S. Ahn, *Food Chem.* **2024**, *439*, 138082.
[19] W. F. Soares, B. D. Chinchin-Piñan, R. M. Silva, J. E. L. Villa, *Food Control* **2024**, *166*, 110754.
[20] M. Lasalvia, V. Capozzi, G. Perna, *Appl. Sci.* **2022**, *12*, 5345.
[21] N. Kumar, A. Bansal, G. S. Sarma, R. K. Rawal, *Talanta* **2014**, *123*, 186.
[22] S. M. Lundberg, S.-I. Lee, *Adv. Neural Inf. Process Syst.* **2017**, *30*, 1.
[23] R. G. Brereton, G. R. Lloyd, *Analyst* **2010**, *135*, 230.
[24] B. Debska, B. Guzowska-Swider, *Anal. Chim. Acta* **2011**, *705*, 283.
[25] S. Albawi, T. A. Mohammed, S. Al-Zawi, *IEEE, Antalya* **2017**, *2017*, 1.
[26] J. Wu, X.-Y. Chen, H. Zhang, L.-D. Xiong, H. Lei, S.-H. Deng, *J. Electron. Sci. Technol.* **2019**, *17*, 26.
[27] R. R. Selvaraju, M. Cogswell, A. Das, R. Vedantam, D. Parikh, D. Batra, *Int. J. Comput. Vis.* **2020**, *128*, 336.

SUPPORTING INFORMATION

Additional supporting information can be found online in the Supporting Information section at the end of this article.

How to cite this article: N. Yeo, J. M. Han, M. G. Kim, J. Y. Kim, H. Cho, S. Y. Lee, J.-H. Auh, B. H. Kim, S. Ahn, *Bull. Korean Chem. Soc.* **2025**, 1. https://doi.org/10.1002/bkcs.70056