Vol. 21 (2026): Proceedings of 2026 International Conference on Green Chemistry and Biomedical Engineering (IC-GCBE2026)
E-ISSN: 3105-0522 P-ISSN: 3105-0514
DOI: https://doi.org/10.70267/IC-GCBE2026.2026.2633

New Chemometrics, Foodomics and A.I. Technology in Food Analysis

Yufei Huang*

Wuhan Polytechnic University, Wuhan, China

**Corresponding author: Yufei Huang.*

Abstract

Against the backdrop of increasingly complex global food supply chains and growing consumer demands for food quality and safety, traditional targeted detection methods can no longer meet the needs of high-throughput, non-destructive and systematic food analysis. The integrated development of foodomics, chemometrics and artificial intelligence (AI) has emerged as a pivotal research direction in this field. This paper aims to systematically sort out the theoretical framework and technical evolution of the three technologies, summarise their integrated application scenarios, analyse core limitations and propose future development directions. Through a systematic literature review, this study elaborates on the underlying principles and technical logic of each technology, and compares their technical boundaries, performance characteristics and complementary relationships in typical food analysis tasks. The results demonstrate that the three technologies have significant theoretical complementarity and have formed mature application paths in food authenticity traceability, quality evaluation and flavour component analysis. Nevertheless, their integrated application is still constrained by bottlenecks, including uneven data quality, limited model generalisation and insufficient technical standardisation. This study clarifies the integrated development logic of relevant technologies and provides valuable references for food analysis research, industrial practice, food safety supervision and the high-quality development of the food industry.

Keywords

foodomics, chemometrics, artificial intelligence, food analysis, machine learning

1. Introduction

Food analysis is a core discipline of food science that underpins consumer food quality and safety and drives the development of the food industry. As the global food supply chain grows increasingly complex and consumer demand for food quality and safety continues to rise, traditional targeted detection methods based on single components and indicators can no longer fully reveal the internal laws of complex food systems, failing to satisfy modern requirements for rapid, non-destructive and high-throughput analysis [1].

Against this background, foodomics has emerged as a comprehensive research paradigm. It applies multi-omics technologies to systematically study food structure and quality attributes at the molecular level, expanding the analytical scope of food analysis. In parallel, data processing methods in food analysis have evolved from traditional chemometrics to modern machine learning and deep learning, advancing food analysis from simple qualitative identification to quantitative prediction, and from single-component analysis to

Copyright © 2026 The Authors. All individual papers remain the copyright of their respective author(s) · Published by Zeus Academic Research Publishing Group

systematic network analysis [2]. Chemometrics features a sound theoretical system and high interpretability, and is widely used in food spectral analysis and authenticity verification. AI technology, by contrast, can automatically extract data features and fit non-linear relationships, enabling deeper mining of omics and image data.

At present, the cross-integration of foodomics, chemometrics and AI has attracted widespread attention in food analysis, with growing research output. However, challenges remain such as uneven data quality, limited model applicability and low industrialisation feasibility. This paper systematically reviews the development and integrated application of the three technologies, clarifies their respective characteristics and limitations, and proposes future development directions.

2. Research Objective

This review aims to provide a structured, theory-oriented analysis of the technical system of foodomics, chemometrics and AI in food analysis, rather than a simple enumeration of application cases. Specifically, the research objectives are threefold:

Research Objective 1: To elaborate on the underlying theoretical principles and technical logic of foodomics, chemometrics and AI, and clarify their respective evolutionary paths in food analysis.

Research Objective 2: To compare the technical boundaries and performance characteristics of the three methods across different food analysis tasks, and reveal their inherent complementary relationship.

Research Objective 3: To summarise the core bottlenecks restricting the integrated application of the three technologies, and propose feasible development directions for industrial implementation.

These objectives serve as the main thread of the following literature review and discussion.

3. Literature Review

3.1 Theoretical Framework and Analytical Principle of Foodomics

Foodomics was defined by Cifuentes [3] as the application of omics technologies to the food field, which systematically studies the molecular composition and variation rules of food to characterise its quality, safety, nutritional and functional properties. The core theoretical advantage of foodomics lies in its non-targeted, high-throughput analytical approach: it can acquire massive molecular information from food systems, identify biomarkers and action mechanisms that are undetectable by traditional methods, and broaden the research dimension of food science.

Technically, foodomics is a multi-disciplinary system covering metabolomics, lipidomics, proteomics and other branches [4]. Metabolomics focuses on the composition and changes of small-molecule metabolites in food, which can directly reflect food quality and safety status. Lipidomics conducts systematic analysis of lipid components, and is applied to oil quality evaluation and adulteration identification. Proteomics studies the food proteome, and is commonly used for allergen detection and processing characteristic analysis. These branches characterise food properties at different molecular levels, jointly constructing a comprehensive analytical system.

In practice, the implementation of foodomics relies on a complete technical chain, including sample pretreatment, separation and detection, and data analysis [5]. Advances in extraction and purification technologies have improved the extraction efficiency of target components from complex food matrices. Chromatography-mass spectrometry and high-resolution mass spectrometry enable the simultaneous separation and identification of multiple compounds, generating wide-coverage molecular fingerprint data of food. High-resolution mass spectrometry, as the core analytical tool, supports simultaneous qualitative and quantitative analysis of hundreds to thousands of compounds in food, providing a solid data basis for omics research [6].

As a key development direction, multi-omics integration can verify analytical results from different molecular levels and fully reveal the internal mechanism of quality change or safety risk. For example, the combination of metabolomics and lipidomics enables comprehensive analysis of both polar and non-polar

Copyright © 2026 The Authors. All individual papers remain the copyright of their respective author(s) · Published by Zeus Academic Research Publishing Group

compounds in processed and stored food; the combination of proteomics and metabolomics can explore flavour formation mechanisms at both enzyme and metabolite levels [7]. In addition, the research scope of foodomics is extending to nutrition and health. By correlating foodomics data with human nutrition and metabolism data, researchers can explore the mechanism of food components on human health, and optimise dietary plans and nutritional products for different populations [8].

3.2 Mathematical Mechanism of Chemometrics and Spectral Calibration Theory

Chemometrics is a data analysis method system based on multivariate statistics, which has developed over decades into a complete theoretical system widely used in food authenticity identification [9, 10], quality evaluation and component quantification. It has promoted food analysis from simple "instrument detection" to the full process of "instrument detection – data modelling – result discrimination", improving analytical efficiency and data value [11].

From a theoretical perspective, chemometric methods for food analysis can be divided into three main categories, each with clear mathematical logic:

Represented by Principal Component Analysis (PCA), it transforms high-dimensional data into a small number of orthogonal principal components through linear transformation, while retaining the maximum variance of the original data. It is used for data visualisation, cluster analysis and outlier detection, and does not require prior sample labelling [12].

Represented by Partial Least Squares Discriminant Analysis (PLS-DA) and Linear Discriminant Analysis (LDA), it constructs a linear association model between data features and sample categories to realise classification of unknown samples. Such methods require labelled training data and feature clear interpretability [12].

Represented by Partial Least Squares Regression (PLSR), it establishes a quantitative calibration model between analytical signals and food quality indicators, solving the problem of multicollinearity among independent variables. It enables rapid non-destructive determination of target components without complex sample preparation [13].

The combination of spectroscopy and chemometrics is a classic technical route for rapid non-destructive food analysis. Spectroscopic technologies such as infrared spectroscopy, Raman spectroscopy and hyperspectral imaging can rapidly obtain molecular vibration or optical properties of food, with the advantages of high speed, no pretreatment and non-destruction. However, raw spectral data has high dimensionality and overlapping information, so chemometric methods are required for feature extraction and data processing [14]. With the development of imaging technology, hyperspectral imaging can obtain both spatial image information and spectral information of each pixel, which can not only quantify component content but also visualise its spatial distribution to evaluate food quality uniformity [15].

It should be noted that traditional chemometrics has inherent limitations. Its feature extraction relies on manual design, and the model expression ability is limited, making it difficult to fit complex non-linear relationships in food systems. In comparative experiments on complex imaging spectral data, the prediction accuracy and anti-interference ability of traditional chemometric methods are generally lower than those of machine learning methods, especially when the sample matrix is complex and there are many interference factors [16].

Nevertheless, chemometrics is still irreplaceable in scenarios with small sample sizes and high requirements for result interpretability, and the integration of its feature engineering methods with AI modelling has become an important development direction.

3.3 Artificial Intelligence: Nonlinear Feature Learning and Data Mining Logic

AI technology represented by machine learning and deep learning can automatically extract data features and fit non-linear relationships, which has become a key driving force for the intelligent development of food analysis. Different from traditional chemometric methods, deep learning models do not require manual feature selection, and can directly process raw data to extract hidden information, improving the performance of analytical models [17].

Copyright © 2026 The Authors. All individual papers remain the copyright of their respective author(s) · Published by Zeus Academic Research Publishing Group

In food image recognition, Convolutional Neural Networks (CNNs) are the most widely used deep learning architecture. Theoretically, CNNs automatically extract multi-level visual features such as texture, colour and shape from food images through convolutional layers, pooling layers and fully connected layers, realising end-to-end classification, quality assessment, defect detection and foreign body recognition. This technology can be applied to online detection of food processing lines, agricultural product grading and quality monitoring in circulation links [18].

In spectral data analysis, deep learning can directly learn effective features from raw spectra, solving problems such as serious spectral interference and difficult feature extraction under complex matrices, thus improving the accuracy and stability of predictive models. Relevant studies have developed portable near-infrared analysis systems based on deep learning, which realise in-situ rapid detection of mixed powdered food and expand the scope of traditional laboratory analysis [19]. High-throughput optical sensors combined with machine learning can also quickly identify multiple indicators of massive samples, suitable for rapid screening of large quantities of food samples [20].

In food fraud identification and quality control, AI constructs discriminant or prediction models based on food molecular fingerprints, image features or multi-source supply chain data, providing technical support for adulteration detection and dynamic quality management. For example, machine learning can integrate environmental temperature and humidity, logistics time and food quality indices to build a dynamic prediction model for fresh food quality, enhancing quality monitoring and early warning throughout the supply chain [21]. Deep learning can also analyse time-series data of food quality changes and build shelf-life prediction models considering various environmental factors, with higher accuracy than traditional kinetic model methods [22]. In recent years, explainable AI (XAI) has been introduced to solve the "black box" problem of deep learning models, visualising the feature basis of model decisions and improving the reliability of results in regulatory scenarios [23].

In omics data analysis, machine learning provides new tools for mining high-dimensional foodomics data. It can identify key indicators, explore inter-component connections and summarise internal variation rules of food systems, promoting food analysis from descriptive research to mechanistic research [24]. In flavour research, AI models can analyse the non-linear relationship between flavour component composition and sensory perception, predict flavour characteristics of different formulations, and drive flavour R&D from empirical trial and error to data-driven mode [25].

Despite these advances, the application of AI in food science still faces barriers. On one hand, public food datasets have problems such as small quantity, limited scale, non-standard annotation and insufficient diversity, which restrict model generalisation [26]. On the other hand, some studies overemphasise algorithm performance while ignoring food domain knowledge, resulting in models with low practical application value [27]. Standardised operation guidelines for AI modelling in food research are still needed to promote standardised application of related technologies [28].

4. Key Findings

Based on the systematic review of the above literature, three core findings can be summarised:

First, foodomics, chemometrics and AI present significant theoretical complementarity, which is the fundamental driving force for their integrated development. Foodomics provides a systematic molecular data foundation, breaking the limitation of traditional targeted detection in analytical scope; chemometrics guarantees model interpretability and stability under small sample conditions, with mature theoretical support; AI breaks through the bottleneck of complex non-linear relationship fitting, and can automatically mine deep features from high-dimensional data. None of the three technologies can independently cover all demands of modern food analysis, and their integration can realise complementary advantages across the whole chain from molecular characterisation to intelligent decision-making.

Second, there is obvious task adaptability among different technical methods, and no single method is universally optimal. For authenticity identification and classification tasks, machine learning and deep learning models generally show higher prediction accuracy and anti-interference ability than traditional chemometrics. For quantitative calibration tasks with clear physical meaning, chemometric methods such as PLSR have better stability, interpretability and lower sample size requirements. For complex system analysis, such as flavour

Copyright © 2026 The Authors. All individual papers remain the copyright of their respective author(s) · Published by Zeus Academic Research Publishing Group

formation mechanisms and quality change regulation, the combination of multi-omics technology and AI data mining is the most effective technical path.

Third, the core bottleneck restricting the industrial application of related technologies lies in the standardisation system, rather than algorithm iteration. At present, the lack of unified data standards, poor data compatibility between different instruments, and insufficient scale and diversity of public datasets lead to limited model generalisation ability and difficulty in cross-scenario migration. Simply pursuing algorithm complexity cannot solve practical industrial problems; the combination of food domain knowledge, unified data standards and robust evaluation systems is the key to promoting the stable development of AI in the food industry [29].

5. Discussion

5.1 Integrated Application Scenarios of Three Technologies

The integration of foodomics, chemometrics and AI has formed mature technical routes in three core application scenarios:

Omics or spectroscopic technology is used to obtain molecular fingerprint data of food, and then a discriminant model is constructed, combined with chemometrics or AI algorithms to identify food authenticity and origin. For example, untargeted metabolomics combined with chemometrics has been applied to detect non-halal meat adulteration in beef, achieving accurate detection results with good interpretability [30]. At present, mass spectrometry-based omics technology combined with chemometrics and machine learning has become one of the main technical paths for food authenticity identification and quality assessment [31].

Spectroscopy and imaging technology, combined with chemometrics, realises non-destructive determination of food quality characteristics and obtains real-time quality data. On this basis, machine learning models combined with environmental information can predict food quality changes and remaining shelf life, supporting food storage, transportation and sales. This scheme is highly practical for perishable foods such as fresh fruits and vegetables, meat products and dairy products, and can realise dynamic quality control at all stages of the supply chain [32].

Multi-omics technology obtains comprehensive molecular information of flavour-related substances, and AI data mining reveals the molecular regulation mechanism of flavour formation and change, providing a basis for food flavour control and new product development. This promotes flavour research from empirical intuition to data-driven direction [33].

5.2 Existing Technical and Industrial Bottlenecks

Despite the remarkable progress, the integrated application of the three technologies still faces multiple bottlenecks:

First, data quality and standardisation need to be improved. The lack of unified data acquisition standards and annotation norms leads to poor comparability of data from different sources, and the scale and diversity of public datasets are insufficient, which limits the training and generalisation of high-performance models [34].

Second, the model's generalisation ability is limited. Most models are developed for specific food matrices and specific instruments, with poor cross-matrix and cross-instrument migration ability, making it difficult to adapt to complex and changeable actual detection scenarios [35].

Third, technical standardisation is insufficient. There is no unified standard for detection methods and model evaluation systems, and the high cost of large-scale omics and AI detection equipment restricts the popularisation and application of related technologies in primary supervision and small and medium-sized enterprises.

Copyright © 2026 The Authors. All individual papers remain the copyright of their respective author(s) · Published by Zeus Academic Research Publishing Group

5.3 Future Development Directions

In view of the above bottlenecks, future research and industrial practice can be promoted in the following directions: Promote the formulation of unified data acquisition and annotation standards, build large-scale, high-quality public food datasets covering multiple food categories and detection scenarios, and lay a data foundation for model training and generalisation. Integrate food science domain knowledge into AI model construction, develop interpretable fusion models combining chemometric feature engineering and AI modelling, and balance prediction performance and result interpretability. Promote the miniaturisation and low cost of detection equipment combined with AI algorithms, develop portable rapid detection systems suitable for on-site scenarios such as market supervision and production lines, and expand the application scope of related technologies. Formulate unified technical specifications and evaluation standards for detection methods, model construction and result verification, and promote the standardised application of related technologies in the industry.

6. Conclusion

The cross-integration of foodomics, chemometrics and AI is driving the transformation and upgrading of food analysis technology. Foodomics expands the research scope of food analysis and provides a systematic perspective for studying complex food systems; chemometrics, as a mature data analysis method, is a reliable tool for rapid food analysis with stable results and good interpretability; AI technology provides new means for in-depth analysis of complex data and promotes the intelligent development of food analysis. The combination of the three technologies has achieved fruitful results in food authenticity identification, quality evaluation and flavour analysis, providing technical support for the development of the food industry and food safety supervision.

Meanwhile, there are still problems, such as imperfect data standard system, limited model generalisation ability, high technical application cost and insufficient standard specifications. Future research should focus on data standardisation construction, algorithm optimisation, equipment miniaturisation and detection method standardisation, so as to promote the deep integration and practical application of the three technologies, continuously improve the technical level of food analysis, guarantee food safety and facilitate the high-quality development of the food industry.

References

[1] Class, L.-C., Kuhnen, G., Rohn, S., & Kuballa, J. (2021). Diving deep into the data: A review of deep learning approaches and potential applications in foodomics. Foods, 10(8), 1803. https://doi.org/10.3390/foods10081803

[2] Rocchetti, G., Montesano, D., Bordoni, A., & Capozzi, F. (2025). AI-powered advances in data handling for enhanced food analysis: From chemometrics to machine learning. Foods, 14(19), 3415. https://doi.org/10.3390/foods14193415

[3] Cifuentes, A. (2022). Foodomics: Analytical opportunities and challenges. Journal of Agricultural and Food Chemistry, 70(11), 3357–3367. https://doi.org/10.1021/acs.jafc.1c08018

[4] Zhang, S., Chen, J., Gao, F., et al. (2024). Foodomics as a tool for evaluating food authenticity and safety from field to table: A review. Foods, 14(1), 15. https://doi.org/10.3390/foods14010015

[5] Garcia, M. et al. (2025). Techniques used in foodomics. Food Science & Nutrition.

[6] Fraga, C. G. et al. (2024). Foodomics: A sustainable approach for the specific nutrition and diets for human health. Food Science & Nutrition.

[7] Costa, C. et al. (2026). Sensors and mass spectrometry connection for food analysis: A systematic review of methodological synergies. Chemosensors.

[8] Barabasi, A. L. et al. (2026). Quantitative food compounds enable dietary ontology referencing across 500 foods and human plasma. Analytical Chemistry.

Copyright © 2026 The Authors. All individual papers remain the copyright of their respective author(s) · Published by Zeus Academic Research Publishing Group

[9] Djoumbou-Feunang, Y. et al. (2024). Chemical multiverse and diversity of food chemicals. Journal of Chemical Information and Modeling.

[10] Taylor, A. J. et al. (2023). Data-driven elucidation of flavor chemistry. Journal of Agricultural and Food Chemistry.

[11] Rodriguez, J. et al. (2022). An overview on the application of chemometrics tools in food authenticity and traceability. Foods.

[12] Omar, M. et al. (2025). Integration of spectroscopy and chemometric analysis for food authentication: A review. International Food Research Journal.

[13] Li, Y. et al. (2023). Current application of advancing spectroscopy techniques in food analysis: Data handling with chemometric approaches. Foods.

[14] Wang, X. et al. (2025). A comprehensive review on minimally destructive quality and safety assessment of agri-food products: Chemometrics-coupled mid-infrared spectroscopy. Foods.

[15] Liu, H. et al. (2024). Rapid and nondestructive determination of oil content and distribution of potato chips using hyperspectral imaging and chemometrics. ACS Food Science & Technology.

[16] Kelly, S. et al. (2022). Food authentication of almonds (Prunus dulcis Mill.). Fast origin analysis with laser ablation inductively coupled plasma mass spectrometry and chemometrics. Journal of Agricultural and Food Chemistry.

[17] Diaconu, A. et al. (2025). Quantification of paprika adulteration using visible/near-infrared spectroscopy and machine learning. Journal of Food Science and Technology.

[18] Almeida, M. et al. (2025). Rapid authentication of plant-based milk alternatives by coupling portable Raman spectroscopy with machine learning. Journal of AOAC INTERNATIONAL.

[19] Deng, W. et al. (2022). UV fingerprinting approaches for quality control analyses of food and functional food coupled to chemometrics. Molecules.

[20] Loutfi, A. et al. (2023). Review on sensor array-based analytical technologies for quality control of food and beverages. Sensors.

[21] Alkanan, Z. T., Altemimi, A. B., & Hashemi, S. M. B. (2025). Exploring the frontiers of food science: A comprehensive review of advanced magnetic resonance applications in food analysis. Food Science & Nutrition. https://doi.org/10.1002/fsn3.70643

[22] Nicolaï, B. et al. (2025). Comparison between chemometrics and machine learning for the prediction of macronutrients in cheese using imaging spectroscopy. Journal of Food Engineering.

[23] Zhou, W. et al. (2026). Deep learning-based image recognition for food science and technology: End-to-end workflows and domain-specific solutions. Comprehensive Reviews in Food Science and Food Safety.

[24] Zhang, L. et al. (2021). A portable NIR-system for mixture powdery food analysis using deep learning. Infrared Physics & Technology.

[25] Wang, Q. et al. (2026). A review of high-throughput optical sensors for food detection based on machine learning. Sensors.

[26] Moor, C. et al. (2026). The application of artificial intelligence in food fraud and adulteration analysis: A review. Critical Reviews in Food Science and Nutrition.

[27] Spink, J. et al. (2026). The role of artificial intelligence in combating food fraud: A systematic literature review. Critical Reviews in Food Science and Nutrition.

[28] Stella, S. et al. (2025). Revolutionizing fresh food quality control in supply chains with machine learning: Current advances and future challenges. Comprehensive Reviews in Food Science and Food Safety.

[29] Kurek, M. et al. (2026). Deep learning models for shelf-life prediction and regulation of various foods: A systematic review. Journal of Food Science.

Copyright © 2026 The Authors. All individual papers remain the copyright of their respective author(s) · Published by Zeus Academic Research Publishing Group

[30] Kaur, R. et al. (2025). An XAI-enabled 2D-CNN model for non-destructive detection of natural adulterants in the wonder hot variety of red chilli powder. Sustainable Food Technology.

[31] Singh, P. et al. (2026). Rapid detection of milk adulteration using AI-driven portable colorimetric spectroscopy. Scientific Reports.

[32] Ke, Q., Zhang, J., Huang, X., Kou, X., & Zhang, D. (2026). Machine learning unveils three layers of food complexity. npj Science of Food.

[33] Zhang, H. et al. (2025). Artificial intelligence and food flavor: How AI models are shaping the future and revolutionary technologies for flavor food development. Comprehensive Reviews in Food Science and Food Safety.

[34] Van der Putten, P. et al. (2026). Practical guide for food scientists to build AI: Data, algorithms, and applications. Food Chemistry.

[35] Wolfert, J. et al. (2025). Domain knowledge, just evaluation, and robust data standards are required to advance AI in food science. Trends in Food Science & Technology.

Funding

This research received no external funding.

Conflicts of Interest

The authors declare no conflict of interest.

Acknowledgment

This paper is an output of the science project.

Open Access

This chapter is licensed under the terms of the Creative Commons Attribution-NonCommercial 4.0 International License (http://creativecommons.org/licenses/by-nc/4.0/), which permits any noncommercial use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license and indicate if changes were made.

The images or other third party material in this chapter are included in the chapter's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the chapter's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder.



Copyright © 2026 The Authors. All individual papers remain the copyright of their respective author(s) · Published by Zeus Academic Research Publishing Group