Integration of machine learning technologies in food flavor research: Current applications, challenges, and future perspectives

Jing Ling[1*], Dennis R. Heldman[2]

(1. *School of Computer Science, Guangzhou Maritime University, Guangzhou 510725, China*;
2. *Department of Food Science and Technology, The Ohio State University, Columbus, OH 43210, USA*)

Abstract: Flavor characteristics significantly influence consumer food preferences and purchasing behaviors, constituting a vital research domain in food science. While traditional flavor assessment approaches rely primarily on sensory evaluations and instrumental analyses, they face inherent limitations in processing large-scale datasets and generating comprehensive insights. Advanced machine learning (ML) models have revolutionized flavor research through their high-precision predictive capabilities, effectively addressing conventional methodological constraints. These computational approaches enable sophisticated and efficient flavor analysis by integrating multiple data dimensions, including chemical composition (volatile and non-volatile compounds), sensory attributes (taste, aroma, texture), temporal dynamics (flavor release patterns), and consumer responses. ML models demonstrate remarkable capability in simultaneously processing diverse data types, such as gas chromatography-mass spectrometry results, sensory panel evaluations, and real-time flavor release measurements, to predict consumer preferences and optimize flavor formulations. This review examines state-of-the-art ML applications in flavor science, emphasizing crucial areas such as flavor database development, intelligent sensory detection, and food traceability. Through systematic analysis of contemporary ML algorithms, this study critically evaluates their potential and limitations in decoding complex flavor dynamics, providing valuable insights for both researchers and industry practitioners while identifying promising directions for future technological innovations in food flavor analysis and prediction. The comprehensive synthesis presented here represents a significant contribution to the field by establishing a theoretical framework for ML-driven flavor research and offering practical guidelines for the implementation of computational approaches in food flavor analysis.

Keywords: machine learning, food flavor, prediction, flavor perception
DOI: 10.25165/j.ijabe.20251803.9623

Citation: Ling J, Heldman Dennis R. Integration of machine learning technologies in food flavor research: Current applications, challenges, and future perspectives. Int J Agric & Biol Eng, 2025; 18(3): 1–11.

1 Introduction

Food flavor serves as a multifaceted indicator, encompassing critical attributes like product categorization, regional provenance, developmental stage, preservation status, and organoleptic characteristics[1]. These nuanced flavor dimensions fundamentally shape consumer selection and dietary behaviors. Consumer demand for sophisticated flavor profiles has driven significant advances in food science research. Analysis of Scopus-indexed publications in flavor chemistry reveals exponential growth from the 1980s to 2019, with annual publications reaching approximately 700[2], reflecting the field's increasing research emphasis and underscoring the field's expanding research focus.

There are approximately 2500 naturally occurring flavor compounds, many of which can be synthesized. The field of food flavor encompasses extensive data[3], including information on flavor compounds, their interactions, sensory evaluations, consumer preferences, and food formulations[4]. These datasets exhibit nonlinear characteristics due to the complex relationships among various factors influencing flavor perception. The intricate interactions between flavor compounds and the multifaceted nature of sensory perception, influenced by individual differences and environmental factors, further contribute to this nonlinearity[5]. A compelling illustration of such nonlinearity is manifested in the mechanism by which glycerol affects the release and perception of key flavor compounds in baijiu[6], as well as in the interactions between glycerol and specific flavor compounds. Glycerol can selectively enhance the release of certain flavor compounds (e.g., dimethyl trisulfide) while inhibiting the volatility of others (e.g., ethyl acetate) and altering their olfactory detection thresholds. This bidirectional regulatory effect exhibits nonlinear behavior that cannot be predicted by simple linear relationships.

The journey of flavor analysis in food science encompasses four developmental stages, beginning with sensory evaluation methods, progressing through instrumental detection, incorporating integrated analytical approaches, and culminating in machine learning-enhanced automated procedures[7,8]. Initially, sensory analysis provided a comprehensive flavor profile through human evaluation, offering an integrated assessment of food characteristics. However, this method suffers from inherent limitations, primarily subjective biases, and the requirement of extensive evaluator training[9]. To address these challenges, instrumental analysis emerged, employing sophisticated instruments to quantify odor and taste compounds with greater objectivity. Instrumental techniques advanced the field by delivering qualitative and quantitative measurements of flavor compounds. Nevertheless, these methods presented their own constraints, including time-intensive procedures, high costs, and challenges in managing large, complex datasets[10]. Recognizing these limitations, researchers began developing integrated approaches that combined sensory and

Received date: 2024-12-31 **Accepted date:** 2025-03-25
Biographies: Dennis R. Heldman, Professor, research interest: the application of engineering principles and concepts to the manufacturing of foods with minimum energy, water and waste, Email: heldman.20@osu.edu
***Corresponding author: Jing Ling**, Associate professor, research interest: intelligent detection and intelligent information processing. School of Computer Science, Guangzhou Maritime University, Guangzhou 510725, China. Tel: +86-15209508892, E-mail: lingjing0519@163.com.

instrumental methods to leverage the strengths of both techniques. The most recent paradigm shift has been the introduction of machine learning (ML) models in sensory analysis[5]. These intelligent techniques represent a significant leap forward in flavor prediction and adjustment. Unlike traditional detection methods, ML algorithms excel at uncovering latent, previously unknown connections within extensive datasets[11]. By efficiently processing complex data, these models can extract nuanced patterns, correlations, and insights that were previously undetectable. To ensure the reliability and objectivity of ML-based flavor analysis, researchers have developed comprehensive evaluation metrics. These include accuracy, precision, recall, error rate, robustness, and computational complexity, providing a multifaceted framework for assessing the predictive capabilities of[12]. This evolutionary progression demonstrates the continuous refinement of food flavor analysis techniques, highlighting the field's ongoing commitment to more precise, efficient, and insightful methodologies.

The strengths of machine learning lie in its advanced capabilities: handling extensive datasets, detecting intricate relationships within multidimensional spaces, executing self-directed learning from existing information, delivering precise forecasts for novel inputs, and performing swift recognition tasks through self-optimizing computational processes[5]. For instance, strategies such as manipulating microbial communities during wine fermentation to achieve desired flavors[13], improving processing technologies for juices to enhance freshness[14], and increasing trace elements during vegetable cultivation to alter quality and flavor exemplify the application of these techniques[15].

This scholarly review critically examines the recent proliferation of machine learning (ML) techniques in food flavor research, offering a comprehensive analysis of contemporary methodological innovations. The manuscript provides an in-depth exploration of ML implementation processes, performance evaluations, and critical challenges encountered in the field.

The review systematically investigates pivotal research domains, including: flavor database construction, intelligent sensory detection methodologies, and advanced food traceability techniques.

By critically analyzing the current technological landscape, the study aims to:

1) Elucidate emerging ML methodological trends;

2) Identify potential research trajectories;

3) Provide a robust theoretical framework for future food flavor analysis and predictive research.

2 Fundamental principles and techniques of machine learning

Different machine learning (ML) methods are instrumental in extracting valuable insights from complex flavor data, establishing ML as a critical tool in the modern food industry for product development, quality control, and sensory evaluation. By leveraging ML techniques, researchers can gain a deeper understanding of consumer preferences and develop targeted flavor profiles. Notably, support vector machines (SVM), random forests (RF), decision trees (DT), and artificial neural networks (ANN) have emerged as key data processing tools and focal points of research in the field of food flavor analysis[16]. The flowchart depicted in Figure 1 outlines the general process of applying machine learning (ML) methods in food flavor analysis and prediction, which can be categorized into four main stages[17]: data collection and preprocessing, feature extraction, algorithm design and training, and algorithm validation and prediction. Machine learning has gained widespread application in food flavor prediction and regulatory data analysis due to its self-learning capabilities, adaptability, strong fault tolerance, and robustness in mapping complex dynamic nonlinear structures[18].

Implementing machine learning begins with collecting relevant raw data, typically sourced from advanced analytical techniques such as electronic noses[19], gas chromatography-olfactometry (GC-O), machine vision, near-infrared spectroscopy (NIR), and spectral imaging[5,19]. These datasets form the foundation for model development. Following data collection, preprocessing steps are undertaken, which include data cleaning, dimensionality reduction, and data balancing to ensure consistency and quality. Feature

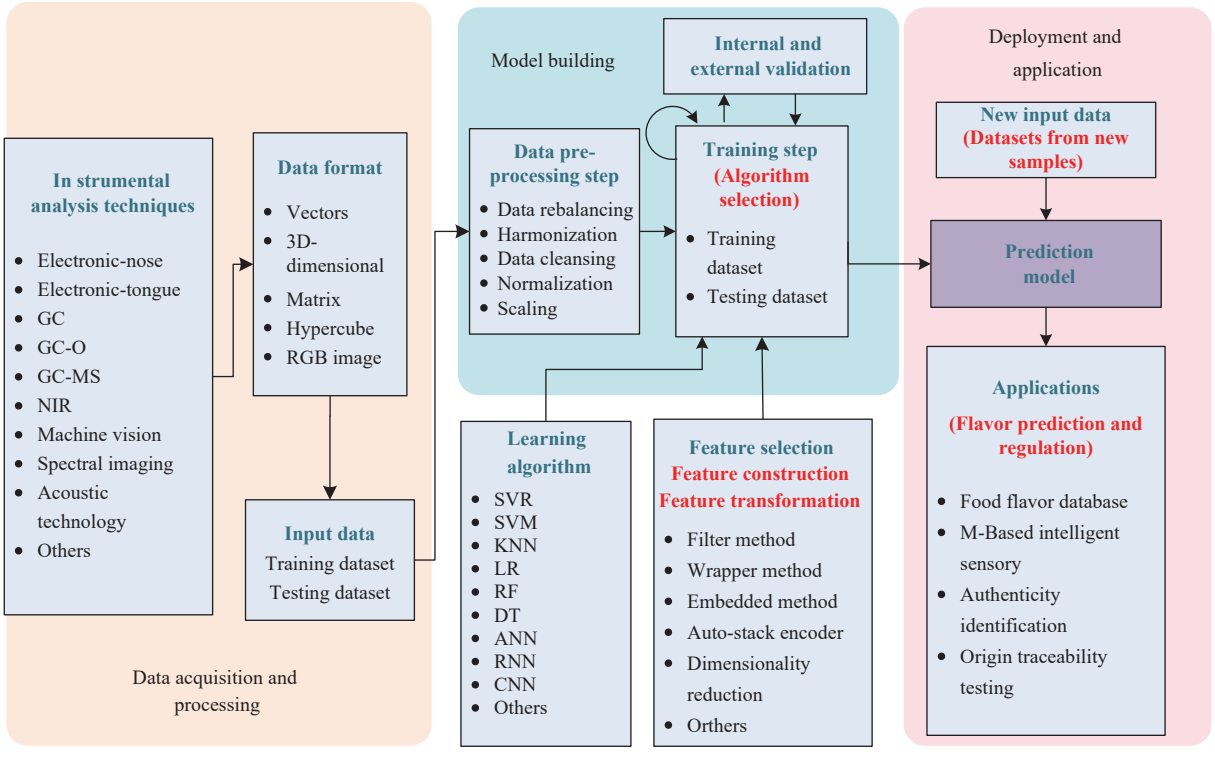


Figure 1 Machine learning workflow for food flavor analysis and prediction

extraction focuses on identifying the most informative variables within the dataset. This process often employs machine learning algorithms such as random forests, autoencoders, and gradient boosting to extract key features that will enhance model performance[20].

During the model training phase, the selected algorithms learn from the processed data to recognize patterns and establish connections between input data and outcome variables. The model's performance is evaluated using a validation set, allowing for further optimization to improve predictive accuracy. The resulting predictive models are then utilized for forecasting new input data outcomes, which may include food flavor quality assessment, freshness detection based on flavor factors, authenticity verification related to food flavor, or traceability of sources.

In summary, machine learning methodologies provide a powerful approach for enhancing food flavor analysis, offering rapid, non-destructive, and reliable quality assessment capabilities. By leveraging comprehensive databases of flavor compounds and sensory profiles, these models can effectively predict various food characteristics including aroma, taste, and textural properties[16].The significant advantages of machine learning have led to its widespread adoption in food quality applications, including honey adulteration detection[21], wine authentication[22], and milk quality evaluation[23]. These applications demonstrate the practical value of machine learning in routine quality control throughout the food production chain.

Machine learning algorithms, integrating supervised, unsupervised, reinforcement learning and deep learning methodologies, have emerged as paradigm-shifting computational strategies in food processing industries, as shown in Figure 2. Supervised learning techniques, including decision trees (DT), support vector machines (SVM), and logistic regression, utilize labeled datasets to create accurate predictive models. These methods effectively map chemical compositions to sensory attributes, enabling precise flavor predictions, quality control enhancements, and detailed flavor profile classifications[24].

Unsupervised Learning methods, including *k*-means clustering and principal component analysis, explore hidden structures within unlabeled data. They excel at revealing intrinsic patterns, uncovering complex relationships, and identifying emerging consumer preferences. Reinforcement Learning optimizes food formulations through iterative interactions with the environment[25]. This approach helps develop adaptive marketing strategies, improve consumer satisfaction, and drive innovative product development. Deep learning models, leveraging neural networks and advanced feature extraction techniques, enable sophisticated flavor prediction and characterization by processing complex multimodal data and identifying intricate relationships between chemical composition, sensory attributes, and consumer perception.

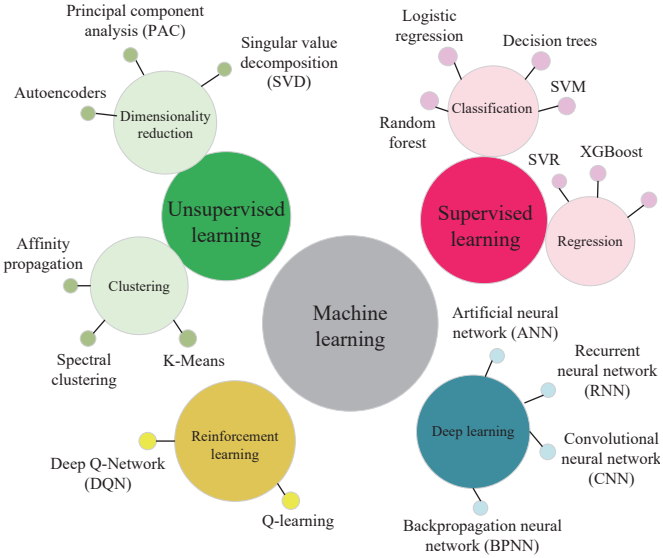
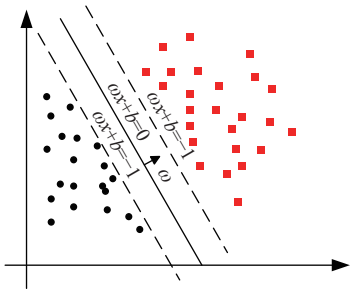


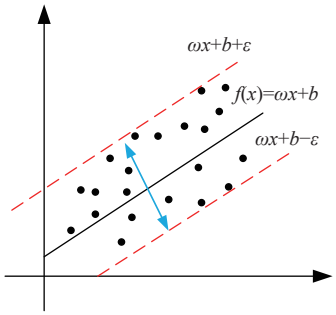
Figure 2 Composition and classification of machine learning

2.1 Traditional machine learning algorithms

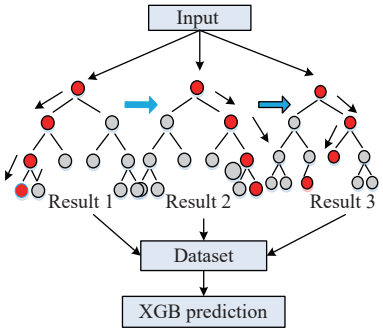
Traditional machine learning algorithms have emerged as powerful analytical tools for food flavor analysis, demonstrating distinct algorithmic advantages in processing complex sensory data



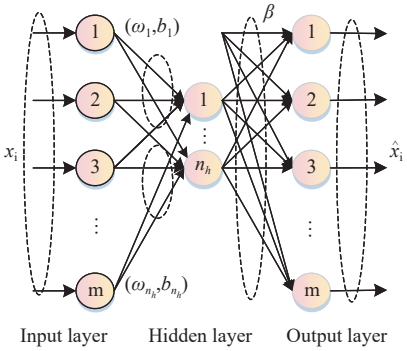
a. Support vector machine (SVM)



b. Support vector regression (SVR)



c. Extreme gradient boosting (XGBoost)



d. Extreme learning machine (ELM)

Figure 3 Schematic diagram of SVM, SVR, XGBoost, and ELM models

(Figure 3). Logistic Regression enables probabilistic modeling of sensory attributes, providing interpretable predictions of flavor characteristics based on processing parameters and ingredient compositions[16]. Support Vector Machines (SVM) and Support Vector Regression (SVR) optimize hyperplane separation in high-dimensional feature spaces, facilitating robust classification and regression of flavor profiles derived from volatile compound data[26]. Gradient Boosting (GB) methods, particularly through ensemble learning and regularization techniques, excel at capturing non-linear interactions between chemical features while mitigating overfitting[27]. Extreme Learning Machines (ELM) offer efficient pattern recognition in spectroscopic and chromatographic data through rapid neural network training[5]. These approaches successfully process multimodal sensor data from electronic noses, mass spectrometry, and spectral imaging for product development applications.

2.1.1 Logistic regression

Given the complex matrix of compounds and physical properties in food systems, univariate regression methods (e.g., Single-Factor Analysis of Variance [ANOVA]) are often insufficient to elucidate the relationships between physicochemical properties, sensory attributes, and distinctive flavors[28]. Several multivariate techniques have been established for analyzing the correlations between flavor compounds and microbial interactions, including Principal Component Analysis (PCA), Partial Least Squares regression (PLS), Ridge regression, and Orthogonal Partial Least Squares regression (OPLS). The selection of regression methods largely depends on the nature of the collected data and computational resource availability. While Logistic regression represents one of the earliest machine learning approaches, its fundamental principles continue to inform and guide the development of contemporary methodologies, particularly in handling binary classification problems in flavor analysis. The key advantages and limitations of these modeling techniques are summarized in Table 1.

Table 1 Recent studies on the application of logistic regression in food flavor analysis

Method	Advantages	Limitations	Sample and measured properties	Reference
Partial Least Squares Discriminant Analysis (PLS)	Suitable for high-dimensional and small sample data; Can build predictive models; Considers relationships between variables.	Model interpretation is relatively poor; May have overfitting risk; Needs appropriate latent variable selection.	**Beer:** Correlate sensory profile and higher alcohol and ester production during fermentation.	[29]
			Liquor: Identification of liquor brands with the same flavor and alcohol content.	[30]
Principal Component Analysis (PCA)	Reduces data dimensionality, simplifies analysis; Extracts main components, reduces noise; Visualizes data, shows sample differences.	May lose some information; Component interpretation is poor; Not suitable for non-linear relationships.	**Craft beers:** Determine the composition and content of amino acids in craft beers.	[31]
			Turkish honeybee: Analyze the organic acids, sugars, minerals, and attributes to antioxidant activities of pollen.	[32]
Orthogonal Partial Least Squares Regression (OPLS)	Separates predictive and non-predictive variables; Improves model prediction and interpretation; Suitable for complex systems.	Computationally intensive; Complex parameter selection; High data requirements.	**Huangjiu:** Analyze the correlation of aroma compounds and identify four major markers (diethyl succinate, furfural, nonanal, and isoamyl alcohol) for aged samples.	[33]
			Yak milk, Investigate the effect of lipoprotein lipase (LPL) on the flavor of yak milk under various storage temperatures.	[34]

2.1.2 SVM and SVR

Support Vector Machine (SVM) and Support Vector Regression (SVR), pioneered by Cortes and Vapnik[35], have evolved into well-established machine learning algorithms widely implemented in food flavor prediction. These algorithms demonstrate exceptional performance in handling nonlinear relationships, small-sample scenarios, and high-dimensional pattern recognition tasks[36].

In food science applications, SVM coupled with sensor arrays has shown remarkable effectiveness in beef flavor prediction. The implementation of radial basis function (RBF) kernels following Principal Component Analysis (PCA) preprocessing has achieved classification accuracies of 90%[37]. However, the inherent sensitivity of SVM/SVR to parameter selection necessitates systematic hyperparameter optimization. Common optimization strategies include grid search with cross-validation for tuning RBF kernel parameters (C and γ), Bayesian optimization for computational efficiency, and evolutionary algorithms for handling complex parameter spaces. Recent studies have explored various optimization strategies to enhance model performance. Han et al. developed a novel non-destructive technique for rice flavor detection by integrating hyperspectral imaging with an improved particle swarm optimization SVM algorithm (GISPSO-SVM), achieving a prediction accuracy of 96%[38]. Wang et al. employed genetic algorithms to optimize SVR parameters for predicting tobacco leaf aroma quality and body, establishing prediction models through variable screening and evaluating performance via modeling validation, leave-one-out cross-validation, and prediction set assessment. The proportion of samples with an absolute error within 0.5 ranged from 84.94% to 97.61%, while the absolute error ranged from 0.16 to 0.27[39].

These advanced optimization techniques have significantly enhanced the practical applicability of SVM/SVR in food flavor prediction, effectively addressing the fundamental challenge of parameter sensitivity while improving model generalization across diverse flavor profiles. The integration of these optimization methods has demonstrated substantial improvements in prediction accuracy across various food matrices, marking a significant advancement in machine learning applications for food flavor analysis.

2.1.3 Gradient Boosting

Gradient Boosting (GB) is an ensemble learning approach that mitigates individual decision trees' limitations by integrating multiple trees to capture complex feature interactions[40]. Specifically, Gradient Boosting Decision Tree (GBDT) iteratively adjusts sample weights to minimize residual errors, progressively enhancing predictive performance. Extreme Gradient Boosting (XGBoost), an advanced GB variant, further optimizes this

approach by minimizing objective functions and improving modeling accuracy[41].

In wine sensory prediction, XGBoost demonstrates remarkable potential by integrating spectral grape extract features with 22 sensory descriptors (color, aroma, flavor, taste, mouthfeel). This methodology successfully predicted 15 sensory descriptors with R^2 values exceeding 0.5, showcasing GB's effectiveness in complex flavor recognition and analysis[42].

2.1.4 Extreme Learning Machine

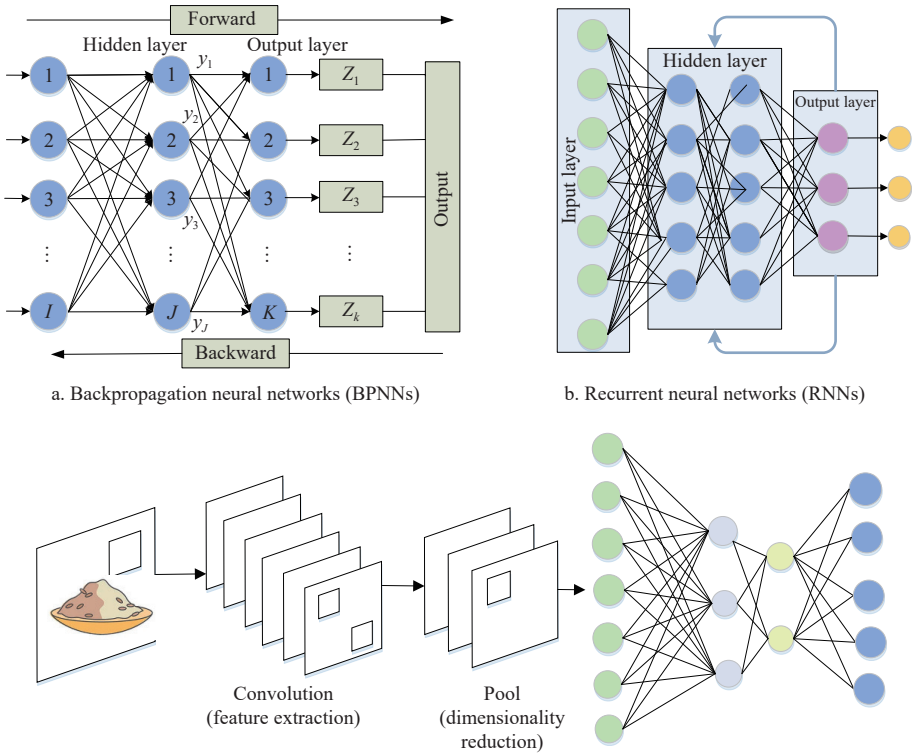
Extreme Learning Machine (ELM) has shown promise in food flavor research due to its simplified training process[43]. By randomly assigning input weights and biases, and subsequently calculating output weights via least-squares, ELM offers computational advantages over traditional backpropagation (BP)[44].

In recent years, ELM has been successfully applied to food flavor classification. For example, in beer flavor prediction using electronic tongue and nose data, ELM achieved 98.33% accuracy, outperforming SVM (96.67%) and RF (94.44%)[45]. ELM also achieved 93.75% accuracy in predicting cocoa bean roasting degree using image-derived color and texture features[46].

While ELM's speed and ease of use are appealing, its reliance on random initialization can lead to performance instability[16]. Ensemble methods, such as bagging, boosting, or random subspace, offer a potential solution to this limitation[47]. For instance, Cheng et al. proposed a multi-disturbance bagging ELM (MdbaggingELM) to improve the robustness and accuracy of the regression model. Their application to cadmium contamination detection in rapeseed achieved a coefficient of determination R^2 of 0.9830 and RMSE of 2.8963 mg/kg[48]. Combining multiple ELM models within an ensemble framework mitigates the variance introduced by random initialization, enhances model robustness, and improves generalization performance, particularly for the complex and high-dimensional flavor datasets common in food science[47].

2.2 Deep learning algorithm

Deep learning (DL) represents a sophisticated machine learning paradigm, leveraging neural network architectures to extract sophisticated features from expansive datasets. As an advanced extension of artificial neural networks, deep learning (DL) has emerged as a robust methodology for addressing highly intricate classification and regression challenges, with prominent algorithms including artificial neural networks, convolutional neural networks (CNNs), recurrent neural networks (RNNs), and backpropagation neural networks (BPNNs). The structural diagrams of these networks are illustrated in Figure 4.



a. Backpropagation neural networks (BPNNs)

b. Recurrent neural networks (RNNs)

c. Convolutional neural networks (CNNs)

Figure 4 Schematic diagram of BPNNs, RNNs, and CNNs models

In the field of food flavor analysis, Traditional Machine Learning (TML) and Deep Learning (DL) architectures exhibit fundamental differences in their computational strategies, as illustrated in Figure 5. DL requires extensive datasets for optimal performance, while TML demonstrates efficacy with limited to moderate data volumes[49]. In order to solve this problem, developing inexpensive parallel computing hardware in the form of graphics processing units (GPUs) is a common way to save model training time for DL algorithms[50].

Training duration represents a significant distinction, with DL frameworks requiring extended periods due to architectural complexity, while TML exhibits shorter training cycles suitable for time-sensitive applications. In task execution, DL excels in concurrent processing, particularly in computer vision applications, achieving superior accuracy in object detection, classification, and segmentation. TML, employing sequential processing, shows optimal performance in focused applications with defined parameters. DL's strengths lie in generalization capabilities and adaptability to novel datasets, while TML offers advantages through reduced computational requirements and expedited training.

Table 2 showcases recent applications of mainstream deep learning algorithms in food flavor analysis. A key challenge in this field has been the traditional reliance on large, labeled sensory datasets. However, emerging research in semi-supervised and self-supervised learning offers promising solutions. Semi-supervised approaches effectively combine limited expert sensory evaluations with abundant unlabeled data, leveraging both taste panel data and unlabeled instrumental measurements, such as electronic nose or mass spectrometry readings. Self-supervised learning provides an alternative by learning representations from unlabeled data through pretext tasks, like predicting masked sensory information or distinguishing variations within food samples[51]. For instance, Wen et al. proposed an active self-semi-supervised learning (AS3L) framework that achieves comparable accuracy to baseline methods in approximately one-third of the training time, demonstrating the potential of these techniques[52].

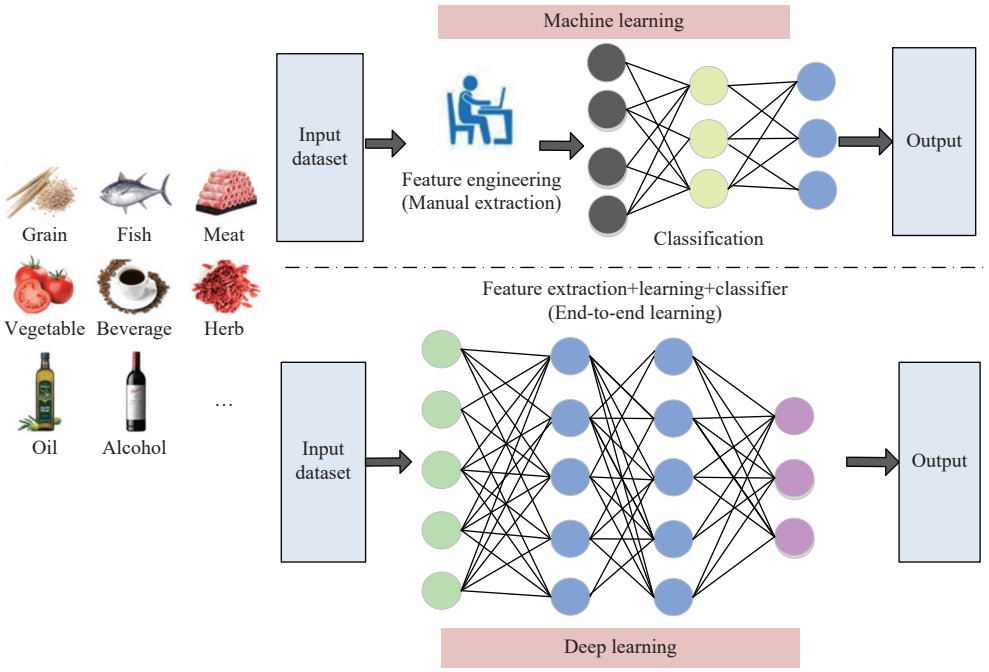


Figure 5 Main structural distinctions between traditional machine learning and deep learning

Table 2 Recent studies on the application of deep learning in food flavor analysis

Method	Advantages	Limitations	Aim of the study	Reference
CNNs	Good at processing image data, can effectively extract spatial features; In flavor component analysis, it can be used to process spectral data, chromatograms, etc., extract features, and perform classification or prediction.	Requires a large amount of labeled data for training; Complex model structure, large amount of calculation.	Predict the olfactory assessment results of peanut oil using the CNN method.	[53]
			A 2D chemical map was used as the input for the CNN model to predict taste (sweet and bitter).	[54]
RNNs	Good at processing sequential data, can capture time dependencies; In flavor analysis, it can be used to process time series data, such as electronic nose and electronic tongue data, and analyze flavor change trends.	Difficult to parallelize, long training time; Prone to vanishing or exploding gradient problems.	The RNN model was established to identify milk adulteration.	[55]
			The RNN model was established to identify adulterated sesame oil.	[56]
BPNNs	Simple model structure, easy to implement; Can be used to process various types of data, with strong versatility.	Easy to fall into local optimal solutions; Requires a large amount of labeled data for training; Poor model interpretability.	Utilizing the BPNN model for pear quality grading, an accuracy rate of 97.4% was achieved.	[57]
			Utilize the Back Propagation Artificial Neural Network (BP-ANN) in conjunction with an electronic nose and GC-IMS to predict the optimal fermentation stage of fermented golden pomfret.	[58]

3 Recent advances of machine learning in food flavor analysis

3.1 Food flavor database utilizing machine learning and big data

The Food Flavor Network, initially developed by Ahn and Ahnert, provided an early platform for data sharing. However, limitations in ingredient coverage highlighted the need for more comprehensive databases. FlavorDB[59], developed by IIT Delhi, addresses this need by cataloging 25 595 spice molecules, encompassing both natural (2254) and synthetic (13 869) compounds[60]. This database integrates molecular characteristics, flavor attributes, and sourcing information, representing a significant expansion in scope.

Complementing FlavorDB are several specialized databases. Food-bridging[61] appears to focus on interactions between food components, although further details are required for a complete assessment. BitterDB[62] specializes in bitter-tasting compounds, providing detailed information on their chemical structures and associated receptors. VirtualTaste[63] employs computational methods to predict taste profiles from molecular structures, offering insights into taste perception.

These databases, while related, serve distinct purposes, as shown in Table 3. FlavorDB's broad coverage of flavor compounds, particularly spices, makes it well-suited for general flavor investigations and the identification of potential flavor combinations[61]. The inclusion of both natural and synthetic molecules is valuable for both culinary innovation and flavor chemistry research. BitterDB's in-depth focus on bitterness is crucial for understanding and modulating this specific taste[62]. It is

an essential resource for research on bitter taste receptors and the development of strategies to reduce bitterness in foods. VirtualTaste's predictive capabilities enable rapid screening of potential flavor compounds, reducing the reliance on extensive sensory evaluation in early-stage flavor development.

The integration of machine learning and big data, exemplified by McCormick's collaboration with IBM[64], further demonstrates the potential of these resources. By analyzing extensive datasets, including flavor profiles and consumer preferences, predictive systems can generate novel flavor combinations, accelerating product development and responding to evolving consumer tastes. These data-driven approaches, in conjunction with the specialized knowledge provided by databases like FlavorDB, BitterDB, and VirtualTaste, are transforming food flavor research and innovation.

Table 3 Comparison of food flavor databases

Database	Data sources	Focus	Features	URL
FlavorDB	"Fenaroli's Handbook of Flavor Ingredients," FooDB	Chemical composition of ingredients; Sensory attributes of compounds	Chemical composition of ingredients; Sensory attributes of compounds; Flavor profiles of food ingredients	https://cosylab.iiitd.edu.in/flavordb2/
BitterDB	"Fenaroli's Handbook of Flavor Ingredients," PubChem, UniProtKB, Merck Index	Bitter taste compounds	Comprehensive collection of bitter compounds; Chemical structures of bitter compounds; Taste properties and associated receptors	http://bitterdb.agri.huji.ac.il/bitterdb/
Virtual Taste	Protein Databases	Taste sensation prediction	Integration of computational models and machine learning algorithms; Prediction and simulation of taste sensations; Chemical properties, sensory data, and human taste responses	https://insilico-cyp.charite.de/VirtualTaste/

3.2 Machine learning-based intelligent sensory technologies

The confluence of analytical instrumentation and machine learning methodologies has transformed modern food flavor analysis. Common instrumental analysis methods include gas chromatography (GC), gas chromatography-olfactometry (GC-O), gas chromatography-mass spectrometry (GC-MS), and high-performance liquid chromatography (HPLC), all of which are utilized to analyze flavor compounds in food[65]. These technologies enable comprehensive characterization of flavor compounds, generating sophisticated chemical signatures encompassing chromatographic profiles, mass spectral data, and quantitative compound distributions, as demonstrated in Table 4.

Table 4 Recent studies on the application of ML combined with instrumental analysis or flavor database in food flavor analysis

Analysis technologies	Methods	Applications	Reference
Electronic nose, electronic tongue	ANN, RF,SVR,XGBoost	Prediction of the freshness of horse mackerel during frozen storag	[66]
Electronic nose	ANN	Prediction of coffee aroma profile and intensity	[67]
	ANN	Prediction of beer aroma	[68]
	CNN	Prediction of odor pleasantness	[69]
GC-MS	CNN, PLSR	Prediction of flavor in peanut oil samples	[53]
	Radio Frequency Regression	Prediction of flavor in sweet peppers at different harvest times	[70]
	ANN	Prediction of wine aroma profile	[71]
GC-IMS	PCA-LDA, PLS-DA, KNN, SVM, XGBoost, ANN	Prediction of sensory quality grading in wine	[72]
GC×GC/TOF-MS	RF, SVM	Prediction of flavor types in Chinese liquor	[73]
GC	ANN, PLS	Prediction of flavor intensity in black currant concentrate	[74]
Near-infrared hyperspectral imaging	PLS, SVM, ELM, CNN	Prediction of off-flavors in farmed salmon	[75]
Raman spectroscopy	SVM	Prediction of flavor in wine	[76]

Notes: The abbreviations are as follows: Support Vector Machine (SVM); Random Forest (RF); Extreme Learning Machine (ELM); Linear Discriminant Analysis (LDA); Back Propagation Artificial Neural Network (BP-ANN); Artificial Neural Network (ANN); Convolutional Neural Network (CNN); Partial Least Squares Regression (PLSR); Principal Component-Linear Discriminant Analysis (PCA-LDA); Partial Least Squares Discriminant Analysis (PLS-DA); k-Nearest Neighbors (KNN); Extreme Gradient Boosting (XGBoost); Partial Least Squares (PLS); Dynamic Programming (DP).

These analytical data undergo preprocessing before integration into machine learning models. The integrated analytical framework encompasses four essential elements: instrumental data acquisition, systematic preprocessing (including signal optimization and data normalization), feature extraction coupled with model development, and analytical task execution for compound identification and prediction. This framework requires sophisticated hardware infrastructure, including high-resolution analytical instruments for detecting trace flavor compounds and high-performance computing systems equipped with GPU accelerators for processing large-scale datasets and complex model architectures[77]. The acceleration in performance compared to traditional central processing units typically ranges from 10 to 40 times, enabling the training of intricate models with up to 10 million parameters to be completed within days rather than weeks or months[78].

A representative example is the electronic nose (e-nose) system, as shown in Figure 6[79], which emulates mammalian olfactoryperception.Thee-nosearchitectureintegratesthreecriticalcomponents: a sensor array (comprising MOS, MOSFET, conductive polymer, and piezoelectric sensors[80]), data preprocessing to remove environmental artifacts, and analytical task execution incorporating pattern recognition algorithms (PCA, PLS, and neural networks) for sample classification and volatile compound identification. This system enables rapid and reproducible analysis of complex flavor profiles in food samples.

These integrated analytical systems demonstrate versatile applications across food science disciplines, particularly in flavor profiling and classification, quality assurance, and geographical origin authentication. Furthermore, these systems facilitate consumer preference analysis through the correlation of analytical data with sensory evaluation, providing valuable insights for product development strategies.

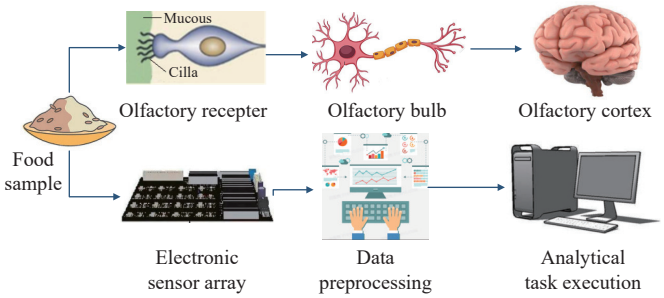


Figure 6 Electronic nose system emulating human olfactory function

3.3 Food authenticity identification and origin traceability testing

The intersection of machine learning algorithms and food flavor analysis has emerged as a promising research domain, particularly in establishing the authenticity and origin of food products. In the domain of beverage authentication, researchers explored the complex volatile compounds in Pinot Noir wines, examining their origins, impact on aroma, and how viticultural, enological, and geographical factors influence the wine's sensory characteristics[81]. Tian et al. investigated meat adulteration using advanced statistical modeling techniques, including Multiple Linear Regression (MLR), Partial Least Squares (PLS), and Least Squares Support Vector Machine (LS-SVM) models. By systematically preparing meat samples with varying proportions of pork/chicken mixed with lamb, they demonstrated the robust predictive capabilities of MLR and PLS regression in determining meat composition[82]. Sabilla et al. advanced the field by integrating deep neural networks (DNN) with electronic nose technologies, successfully differentiating meat types (beef, chicken, and pork) and their anatomical origins (upper body and lower limbs)[83].

These studies collectively underscore the transformative potential of machine learning algorithms in food science, offering unprecedented precision in flavor analysis, quality assessment, and origin verification across diverse food categories.

4 Challenges and future perspectives of machine learning algorithms in food flavor prediction

4.1 Intelligent analysis of consumer feedback

The convergence of machine learning (ML) and advanced big data analytics has revolutionized the approach to understanding consumer flavor perceptions. Hamilton et al. pioneered a breakthrough in sensory attribute analysis through natural language processing (NLP) techniques, effectively addressing the subjectivity inherent in traditional sensory evaluation methods. Utilizing whiskey non-standard text as a case study, the research developed an innovative tool capable of efficiently processing massive volumes of free-text data[84]. Through correspondence analysis and hierarchical clustering, the researchers systematically integrated analysis results into a flavor wheel.

While the ML-driven approach demonstrates significant potential in sensory analysis, several ethical considerations warrant examination. Primary concerns center on data privacy and algorithmic bias[85]. The processing of consumer feedback data necessitates explicit consent protocols and robust anonymization procedures to comply with data protection regulations[86]. Moreover, the ML algorithms may exhibit cultural biases in language processing, potentially skewing sensory descriptions across different cultural contexts.

This study transcends traditional methodological boundaries, demonstrating significant academic and practical value across multiple dimensions, including big data processing, research methodology innovation, and machine learning applications.

4.2 Intelligent ingredient optimization

In the domain of flavor formulation innovation, machine learning has demonstrated exceptional optimization potential. Al-Rifaie et al. employed evolutionary computation techniques to innovatively map brewers' desired sensory characteristics to specific ingredients. By simulating various quantity and combination methods, they precisely designed beer formulations meeting specific sensory requirements[87]. Veeramachaneni et al. further expanded this approach, utilizing machine learning models to predict novel flavor combinations from identical ingredients, successfully gaining validation from sensory evaluation panels[88].

By leveraging machine-learning-assisted algorithms, researchers can elucidate the intricate relationship between molecular chemical properties and the evoked flavor sensations[89]. This approach enables the rational design and in silico prediction of novel flavor molecules, thereby circumventing the need for labor-intensive empirical synthesis.

4.3 Multimodal deep learning for flavor analysis and prediction

Multimodal deep learning represents a cutting-edge technological breakthrough in food science, offering an unprecedented approach to flavor analysis and prediction[90]. This innovative methodology constructs a comprehensive, dynamic flavor perception system by integrating heterogeneous data sources, including spectral information, image data, acoustic features, and chemical composition analyses. Shen et al. employed a suite of advanced sensory technologies, including HS-SPME-GC/MS, e-noses, e-tongues, computer vision, and texture analyzers, to characterize lamb shashliks. They identified key volatile compounds and developed a cross-channel sensory transformer model that accurately predicts sensory attributes across various roasting methods. This research highlights the potential of multimodal deep learning to simulate and enhance sensory evaluation techniques[91].

Unlike traditional single-modal analysis techniques, multimodal deep learning can capture the intricate internal relationships of food flavors, transcending the limitations of single data types[92]. Taking the application of hyperspectral imaging technology in food quality inspection as an example, hyperspectral imaging (HSI) provides more highly localized spectral and spatial information in the image domain as compared to visible/infrared spectroscopy (VIS/IR).

According to most studies utilizing HSI technology, data fusion exists at three distinct levels (see Figure 7)[93,94]: sensor-level fusion, which simply integrates two variables as inputs; feature-level

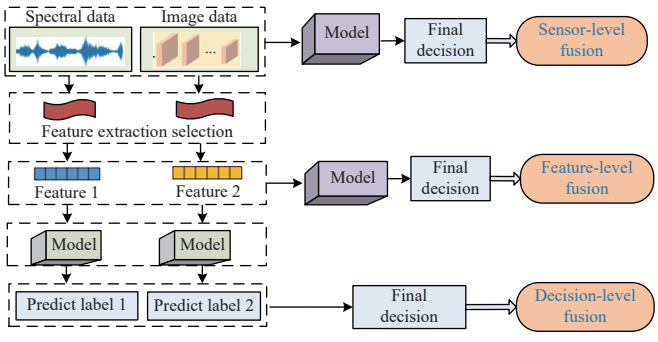


Figure 7 Multimodal data fusion: a general structure diagram

fusion, which combines features after their respective selection; and decision-level fusion, which considers model results built on each data source to make final decisions. The hierarchical data fusion strategy can effectively improve prediction accuracy when applying supervised and unsupervised algorithms for classification when using HSI[94].

5 Conclusions

Despite the considerable potential of machine learning in food flavor prediction, several critical challenges persist. Data heterogeneity and the complexity of cross-modal integration necessitate the development of advanced feature extraction and integration methodologies, alongside rigorous data quality and model reliability assurance. Computational complexity, model interpretability, and algorithmic generalization remain primary bottlenecks. Therefore, future research must prioritize the following five critical domains, ranked according to their urgency in addressing current limitations and their potential to advance the field:

1) Development of robust cross-modal learning algorithms

Food flavor perception is inherently a multi-sensory experience, integrating visual, olfactory, gustatory, and tactile information. Cross-modal learning offers a promising avenue for synthesizing these diverse modalities and exploring their intricate associations, leading to more accurate and nuanced flavor predictions. Effective integration of diverse sensory modalities is foundational for comprehensive prediction systems, making this area the highest priority.

2) Enhancement of model interpretability

Understanding how models extract flavor features from complex data is crucial for validating predictions and elucidating the biological mechanisms of flavor perception. Improved interpretability will foster scientifically sound flavor prediction models and support personalized food development.

3) Establishment of precise flavor component association mechanisms

Accurately defining interactions between flavor components at molecular and perceptual levels is essential for enhancing prediction accuracy. This requires integrating chemical analysis, sensory evaluation, and computational modeling to build comprehensive association models.

4) Development of universal prediction frameworks across food categories

While appealing, the significant variation in flavor profiles across food categories suggests that initially, highly optimized models for specific categories may be more effective. Universal framework development can be progressively pursued in subsequent research.

5) Optimization of computational efficiency

Although important for practical applications, computational efficiency is typically a post-development optimization. Initial research should prioritize model accuracy and interpretability.

Implementation of these research priorities will bridge the gap between computational modeling and practical food industry applications, enhancing product development and consumer experiences. This focused approach will facilitate the transition from theoretical understanding to practical innovation in food science and sensory technologies. The advancement of cross-modal learning algorithms, which underpins all other priorities, is paramount. Effective integration of diverse modalities is critical for accurate and comprehensive flavor prediction, driving innovation in food science and personalized food experiences. This research paradigm has the potential to transform flavor research and provide scientific validation for personalized food development and consumer experience optimization.

Acknowledgements

This work was primarily supported by the National Natural Science Foundation of China (Grant No. 61663039). Additional support was provided by a grant from the Sugar Creek Packing Company and the Dale A. Seiberling Endowment to the Food Engineering Research Laboratory at The Ohio State University.

[References]

[1] Fisher C, Scott T R, Taylor A, Mottram D S. Food flavours: Biology and chemistry. The Royal Society of Chemistry, 1997. 231p. doi: 10.1039/9781847550866.

[2] Queiroz L ., Nogueira I B R, Ribeiro A M. Flavor engineering: A comprehensive review of biological foundations, AI integration, industrial development, and socio-culturaldynamics. Food Research International. 2024; 196: e115100. doi:10.1016/j.foodres.2024.115100.

[3] Sun X X, Yu L L, Xiao M F, Zhang C C, Zhao J X, Narbad A, et al. Exploring core fermentation microorganisms, flavor compounds, and metabolic pathways in fermented Rice and wheat foods. Food Chemistry, 2025; 463: 141019.

[4] Chen X, Zhang W G, Quek S Y, Zhao L Y. Flavor–food ingredient interactions in fortified or reformulated novel food: Binding behaviors, manipulation strategies, sensory impacts, and future trends in delicious and healthy food design. Comprehensive Reviews in Food Science and Food Safety, 2023; 22(5): 4004–4029.

[5] Ji H Z, Pu D D, Yan W J, Zhang Q C, Zuo M, Zhang Y Y. Recent advances and application of machine learning in food flavor prediction and regulation. Trends in FoodScience & Technology, 2023; 138: 738–751.

[6] Wang G N, Liu F F, Pan F, Li H H, Zheng F P, Ye X Q, et al. Study on the interaction between polyol glycerol and flavor compounds of Baijiu: A new perspective of influencing factors of baijiu flavor. Journal of Agricultural and Food Chemistry, 2024; 72(48): 26832–26845.

[7] Wang S Q, Chen H T, Sun B G. Recent progress in food flavor analysis using gas chromatography-ion mobility spectrometry (GC-IMS). Food Chemistry. 2020; 315: 126158. doi: 10.1016/j.foodchem.2019.126158.

[8] Zhang J, Kang D C, Zhang W G, Lorenzo J M. Recent advantage of interactions of protein-flavor in foods: Perspective of theoretical models, protein properties and extrinsic factors. Trends in Food Science & Technology, 2021; 111: 405–425.

[9] Lemarcq V, Van de Walle D, Monterde V, Sioriki E, Dewettinck K. Assessing the flavor of cocoa liquor and chocolate through instrumental and sensory analysis: A critical review. Critical Reviews in Food Science and Nutrition, 2021; 62(20): 5523–5539.

[10] Khan M I H, Sablani S S, Nayak R, Gu Y T. Machine learning‐based modeling in food processing applications: State of the art. Comprehensive Reviews in Food Science and Food Safety, 2022; 21(2): 1409–1438.

[11] Ji X Y, Ji W S, Ding L F. Untargeted GC-MS metabolomics combined with multivariate statistical analysis as an effective method for discriminating the geographical originof shrimp paste. Food Analytical Methods, 2024; 17: 200–206.

[12] Rojas C, Ballabio D, Consonni V, Suárez-Estrella D, Todeschini R. Classification-basedmachine learning approaches to predict the taste of molecules: A review. Food Research International, 2023; 171: 113036.

[13] Barbosa C, Ramalhosa E, Vasconcelos I, Reis M, Mendes-Ferreira A. Machine learning techniques disclose the combined effect of fermentation conditions on yeast mixed-culture dynamics and wine quality. Microorganisms, 2022; 10(1): 107.

[14] Niu H H, Zhang M, Yu Q, Liu Y P. Status and trends of artificial intelligence in the R & D of future fruit & vegetable juice. Innovative Food Science & Emerging Technologies, 2024; 97: 103796.

[15] Pandey V K, Srivastava S, Dash K K, Singh R, Mukarram S A, Kovács B, HarsányiE. Machine learning algorithms and fundamentals as emerging safety tools in preservation of fruits and vegetables: A review. Processes, 2023; 11(6): 1720.

[16] Zeng X Q, Cao R, Xi Y, Li X J, Yu M H, Zhao J L, et al. Food flavor

analysis 4.0: A cross-domain application of machine learning. Trends in Food Science & Technology, 2023; 138: 116–125. doi: 10.1016/j.tifs.2023.06.011.

[17] Ashmore B, Calinescu R, Paterson C. Assuring the machine learning lifecycle: Desiderata, methods, and challenges. ACM Computing Surveys, 2021; 54(5): 1–39.

[18] Greener J G, Kandathil S M, Moffat L, Jones D T. A guide to machine learning for biologists. Nature reviews Molecular Cell Biology, 2022; 23: 40–55.

[19] Liu W, Fan Y, Liu Q N, Xu F, Zhang L, Hu H H. Identification of flavor profiles of Chinese pancakes from various areas using smart instruments combined with E-noses and E-tongues. Int J Agric & Biol Eng, 2023; 16(1): 283–290.

[20] Ao Y, Li H Q, Zhu L P, Ali S, Yang Z G. The linear random forest algorithm and its advantages in machine learning assisted logging regression modeling. Journal of Petroleum Science and Engineering, 2019; 174: 776–789.

[21] Wu L M, Du B, Vander Heyden Y, Chen L Z, Zhao L W, Wang M, et al. Recent advancements in detecting sugar-based adulterants in honey - A challenge. TrAC Trends in Analytical Chemistry. 2017; 86: 25–38.

[22] Ranaweera R K R, Capone D L, Bastian S E P, Cozzolino D, Jeffery D W. A review of wine authentication using spectroscopic approaches in combination with chemometrics. Molecules, 2021; 26(14): 4334.

[23] Nanou E, Pliatsika N, Stefas D, Couris S. Identification of the animal origin of milk via Laser-Induced Breakdown Spectroscopy. Food Control, 2023; 154: 110007.

[24] Naciri L C, Mastinu M, Melis M, Green T, Wolf A, Hummel T, et al. A supervised learning regression method for the analysis of the taste functions of healthy controls and patients with chemosensory loss. Biomedicines, 2023; 11(8): 2133.

[25] Queiroz L P, Rebello C M, Costa E A, Santana V V, Rodrigues B C L, Rodrigues A E, et al. A reinforcement learning framework to discover natural flavor molecules. Foods, 2023; 12(6): 1147.

[26] Zou X G, Xin C R, Wang C Y, Li Y H, Wang S C, Zhang W T, et al. Non-destructive detection of chicken freshness based on multiple features image fusion and support vector machine. Int J Agric & Biol Eng, 2024; 17(6): 264–272.

[27] Yang Z F, Xiao R, Xiong G L, Lin Q L, Liang Y, Zeng W B, et al. A novel multi-layer prediction approach for sweetness evaluation based on systematic machine learning modeling. Food Chemistry, 2022; 372: 131249.

[28] Qiu S Z, Han H Y, Zeng H, Wang B. Machine learning based classification of yogurtaroma types with flavoromics. Food Chemistry, 2024; 438: 138008.

[29] He Y, Dong J J, Yin H, Zhao Y X, Chen R, Wan X J, et al. Wort composition and its impact on the flavour-active higher alcohol and ester formation of beer-a review. Journal of the Institute of Brewing, 2014; 120(3): 157–163.

[30] Yang B, Yao L J, Pan T. Near-infrared spectroscopy combined with partial least squares discriminant analysis applied to identification of liquor brands. Engineering, 2017; 9(2): 181–189.

[31] Aredes R S, Peixoto F C, Sphaier L A, de Carvalho Marques F F. Evaluation of craft beers through the direct determination of amino acids by capillary electrophoresis and principal component analysis. Food Chemistry, 2021; 344: 128572.

[32] Kalaycıoğlu Z, Kaygusuz H, Döker S, Kolaylı S, Erim F B. Characterization of Turkish honeybee pollens by principal component analysis based on their individual organicacids, sugars, minerals, and antioxidant activities. LWT, 2017; 84: 402–408.

[33] Li M, Zhan P, Wang P, Tian H L, Geng J Z, Wang L X. Characterization of aroma-active compounds changes of Xiecun Huangjius with different aging years based on odor activity values and multivariate analysis. Food Chemistry, 2023; 405: 134809.

[34] Shu S B, Jing R, Li L, Wang W H, Zhang J C, Luo Z, et al. Effects of different heat treatments on yak milk proteins on intestinal microbiota and metabolism. Foods, 2024; 13(2): 192.

[35] da Silva Santos C E, Sampaio R C, dos Santos Coelho L, Bestard G A, Llanos C H. Multi-objective adaptive differential evolution for SVM/SVR hyperparameters selection. Pattern Recognition, 2021; 110: 107649.

[36] Feng Y H, Wang Y, Beykal B, Qiao M Y, Xiao Z L, Luo Y C. A mechanistic review on machine learning-supported detection and analysis of volatile organic compounds for food quality and safety. Trends in Food Science & Technology, 2024; 143: 104297.

[37] Wang H M, Wang X D, Xing Y W, Li X, Daun J J. Evaluation of beef flavor attribute based on sensor array in tandem with support vector machines. Food Measure, 2019; 13: 2663–2671.

[38] Yang H, Qu F H, Yang Y, Li X F, Wang P, Guo S K, et al. Study on the determination of flavor value of rice based on grid iterative search swarm optimization support vector machine model and hyperspectral imaging. Sensors, 2024; 24(14): 4635.

[39] Wang J W, Zhang Y L, Liu T A, Lu T, Wu B J, Duan W D. Prediction of flue-cured tobacco leaves aroma quality and volume based on genetic algorithm and support vector machine. In: 2023 2nd International Conference on Artificial Intelligence, Human-Computer Interaction and Robotics (AIHCIR), Tianjin, China: IEEE. 2023; pp.140–145. doi: 10.1109/AIHCIR61661.2023.00031.

[40] Li P, Geng J, Li H C, Niu Z Y. Fish meal freshness detection by GBDT based on a portable electronic nose system and HS-SPME–GC–MS. European Food Research and Technology, 2020; 246: 1129–1140.

[41] Shimizu T, Hattori Y, Ogami S, Imayoshi Y, Moriya S, Okada K, et al. Identification of aroma compounds responsible for apple flavor via palatability score prediction in tree-based regression models. LWT, 2024; 192: 115737.

[42] Armstrong C E J, Niimi J, Boss P K, Pagay V, Jeffery D W. Use of machine learning with fused spectral data for prediction of product sensory characteristics: The case of grape to wine. Foods, 2023; 12(4): 757.

[43] Huang G-B, Zhu Q-Y, Siew C-K. Extreme learning machine: Theory and applications. Neurocomputing, 2006; 70(1-3): 489–501.

[44] Ding S F, Xu X Z, Nie R. Extreme learning machine and its applications. Neural Computing and Applications, 2014; 25: 549–556.

[45] Men H, Shi Y, Fu S L, Jiao Y N, Qiao Y, Liu J J. Mining feature of data fusion in the classification of beer flavor information using E-Tongue and E-Nose. Sensors, 2017; 17(7): 1656.

[46] Yang Y, Darwish A G, El-Sharkawy I, Zhu Q B, Sun S P, Tan J Z. Rapid determination of the roasting degree of cocoa beans by extreme learning machine (ELM)-based imaginganalysis. Journal of Agriculture and Food Research, 2022; 10: 100437.

[47] Liu Y, Cao W P, Ming Z, Wang Q, Zhang J Y, Xu Z W. Ensemble neural networks with random weights for classification problems. In: Proceedings of the 2020 3rd International Conference on Algorithms, Computing and Artificial Intelligence (ACAI '20). New York, NY, USA: Association for Computing Machinery. 2020; pp.1–5. doi: 10.1145/3446132.3446147.

[48] Cheng J H, Sun J, Yao K S, Xu M, Wang S M, Fu L H. Development of multi-disturbance bagging Extreme Learning Machine method for cadmium content prediction of rape leaf using hyperspectral imaging technology. Spectrochimica Acta Part A: Molecular and Biomolecular Spectroscopy, 2022; 279: 121479.

[49] Janiesch C, Zschech P, Heinrich K. Machine learning and deep learning. Electronic Markets, 2021; 31(3): 685–695.

[50] Zhou L, Zhang C, Liu F, Qiu Z, He Y. Application of deep learning in food: a review. Comprehensive Reviews in Food Science and Food Safety, 2019; 18(6): 1793–1811.

[51] Cai Z W, Ravichandran A, Maji S, Fowlkes C, Tu Z, Soatto S. Exponential moving average normalization for self-supervised and semi-supervised learning. In: 2021 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), Nashville, TN, USA: IEEE, 2021; pp.194–203. doi: 10.1109/CVPR46437.2021.00026.

[52] Wen Z T, Pizarro O, Williams S. Active self-semi-supervised learning for few labeled samples. arXiv: 2203.04560, 2024; In press. doi: 10.48550/arXiv.2203.04560.

[53] Bi K X, Zhang D, Qiu T, Huang Y Z. GC-MS fingerprints profiling using machine learning models for food flavor prediction. Processes, 2020; 8(1): 23.

[54] Bo W C, Qin D Y, Zheng X, Wang Y, Ding B T, Li Y H, et al. Prediction of bitterant and sweetener using structure-taste relationship models based on an artificial neural network. Food Research International, 2022; 153: 110974.

[55] Nagamalla V, Kumar B M, Janu N, Preetham A, Gangadharan S M P, Alqahtani M A, et al. Detection of adulteration in food using recurrent neural network with internet of things. Journal of Food Quality, 2022; 2022(1): 6163649.

[56] Wu X J, Zhao Z L, Tian R L, Niu T D, Gao S B, Liu H L. Total synchronous fluorescence spectroscopy coupled with deep learning to rapidly identify the authenticity of sesame oil. Spectrochimica Acta Part A:

Molecular and Biomolecular Spectroscopy, 2021; 244: 118841.
[57] Choi H S, Cho J B, Kim S G, Choi H S. A real-time smart fruit quality grading system classifying by external appearance and internal flavor factors. In: 2018 IEEE International Conference on Industrial Technology (ICIT), Lyon, France: IEEE, 2018; pp.2081–2086. doi: 10.1109/icit.2018.8352510.
[58] Chen Q, Wang Y Q, Wu Y Y, Li C S, Li L H, Yang X Q, et al. Investigation of fermentation-induced changes in the volatile compounds of Trachinotus ovatus (meixiangyu) based on molecular sensory and interpretable machine-learning techniques: Comparison of different fermentation stages. Food Research International, 2021; 150: 110739.
[59] Garg N, Sethupathy A, Tuwani R, NK R, Dokania S, Iyer A, et al. FlavorDB: a database of flavor molecules. Nucleicacids Research, 2018; 46(D1): D1210–D1216.
[60] Kou X R, Shi P Q, Gao C K, Ma P H, Xing H D, Ke Q F, et al. Data-driven elucidation of flavor chemistry. Journal of Agricultural and Food Chemistry, 2023; 71(18): 6789–6802.
[61] Simas T, Ficek M, Diaz-Guilera A, Obrador P, Rodriguez P R. Food-bridging: A new network construction to unveil the principles of cooking. Frontiers in ICT, 2017; 4: 14.
[62] Dagan-Wiener A, Di Pizio A, Nissim I, Bahia M S, Dubovski N, Margulis E, et al. BitterDB: Taste ligands and receptors database in 2019. Nucleic Acids Research, 2019; 47(D1): D1179–D1185.
[63] Fritz F, Preissner R, Banerjee P. VirtualTaste: A web server for the prediction of organoleptic properties of chemical compounds. Nucleic Acids Research, 2021; 49(W1): W679–W684.
[64] Das P, Varshney L R. Explaining artificial intelligence generation and creativity: Human interpretability for novel ideas and artifacts. IEEE Signal Processing Magazine, 2022; 39(4): 85–95.
[65] Benavides J, Rojas J, Lopera Y. Instrumental analysis of flavors and fragrances in food processing: Hyphenated techniques. Flavors and Fragrances in Food Processing: Preparation and Characterization Methods, 2022; 1433: 341–359.
[66] Li H Y, Wang Y, Zhang J X, Li X P, Wang J X, Yi S M, et al. Prediction of the freshness of horse mackerel (Trachurus japonicus) using E-nose, E-tongue, and colorimeter based on biochemical indexes analyzed during frozen storage of whole fish. Food Chemistry, 2023; 402: 134325.
[67] Liu M, Han X M, Tu K, Pan L Q, Tu J, Tan L, et al. Application of electronic nose in Chinese spirits quality control and flavour assessment. Food Control, 2012; 26(2): 564–570.
[68] Gonzalez Viejo C, Fuentes S. Beer aroma and quality traits assessment using artificial intelligence. Fermentation, 2020; 6(2): 56.
[69] Wu D L, Luo D H, Wong K Y, Hung K. POP-CNN: Predicting odor pleasantness with convolutional neuralnetwork. IEEE Sensors Journal, 2019; 19(23): 11337–11345.
[70] Eggink P M, Maliepaard C, Tikunov Y, Haanstra J P W, Pohu-Flament L M M, de Wit-Maljaars S C, et al. Prediction of sweet pepper (*Capsicum annuum*) flavor over different harvests. Euphytica, 2012; 187: 117–131.
[71] Fuentes S, Tongson E, Torrico D D, Viejo C G. Modeling pinot noir aroma profiles based on weather and water management information using machine learning algorithms: A vertical vintage analysis using artificial intelligence. Foods, 2019; 9(1): 33.
[72] Zhu W Y, Benkwitz F, Kilmartin P A. Volatile-based prediction of sauvignon blanc quality gradings with static headspace–gas chromatography–ion mobility spectrometry (SHS–GC–IMS) and interpretable machine learning techniques. Journal of Agricultural and Food Chemistry, 2021; 69(10): 3255–3265.
[73] Li B, Liu M, Lin F, Tai C, Xiong Y F, Ao L, et al. Marker-independent food identification enabled by combing machine learning algorithms with comprehensive GC×GC/TOF-MS. Molecules, 2022; 27(19): 6237.
[74] Boccorh R K, Paterson A. An artificial neural network model for predicting flavour intensity in blackcurrant concentrates. Food Quality and Preference, 2002; 13(2): 117–128.
[75] Sun D W, Zhou C Q, Hu J, Li L, Ye H B. Off-flavor profiling of cultured salmonids using hyperspectral imaging combined with machine learning. Food Chemistry, 2023; 408: 135166.
[76] Leong Y X, Lee Y H, Koh C S L, Phan-Quang G C, Han X, Phang I Y. Surface-enhanced Raman scattering (SERS) taster: A machine-learning-driven multireceptor platform for multiplex profiling of wine flavors. Nano Letters, 2021; 21(6): 2642–2649.
[77] Shahroodi T, Zahedi M, Firtina C, Alser M, Wong S, Mutlu O. Demeter: A fast and energy-efficient food profiler using hyperdimensional computing in memory. IEEE Access, 2022; 10: 82493–82510.
[78] Zhang X L, Yang J, Lin T, Ying Y B. Food and agro-product quality evaluation based on spectroscopy and deep learning: A review. Trends in Food Science & Technology, 2021; 112: 431–441.
[79] Chen H, Huo D X, Zhang J L. Gas recognition in E-nose system: A review. IEEE transactions on biomedical circuits and systems, 2022; 16(2): 169–184.
[80] Panda S, Mehlawat S, Dhariwal N, Kumar A, Sanger A. Comprehensive review on gas sensors: Unveiling recent developments and addressing challenges. Materials Science and Engineering: B, 2024; 308: 117616.
[81] Longo R, Carew A, Sawyer S, Kemp Belinda, Kerslake F. A review on the aroma composition of Vitis vinifera L. Pinot noir wines: Origins and influencing factors. Critical Reviews in Food Science and Nutrition, 2021; 61(10): 1589–1604.
[82] Tian X J, Wang J, Shen R Q, Ma Z R, Li M S. Discrimination of pork/chicken adulteration in minced mutton by electronic taste system. International Journal of Food Science and Technology, 2019; 54(3): 670–678.
[83] Sabilla S I, Sarno R, Triyana K, Hayashi K. Deep learning in a sensor array system based on the distribution of volatile compounds from meat cuts using GC–MS analysis. Sensing and Bio-Sensing Research, 2020; 29: 100371.
[84] Hamilton L M, Lahne J. Fast and automated sensory analysis: Using natural language processing for descriptive lexicon development. Food Quality and Preference, 2020; 83: 103926.
[85] El Mestari S Z, Lenzini G, Demirci H. Preserving data privacy in machine learning systems. Computers & Security, 2024; 137: 103605.
[86] Hu T C, Kyrychenko Y, Rathje S, Collier N, van der Linden S, Roozenbeek J. Generative language models exhibit social identity biases. Nature Computational Science, 2025; 5(1): 65–75.
[87] Al-Rifaie M M, Cavazza M. Evolutionary optimisation of beer organoleptic properties: A simulation framework. Foods, 2022; 11(3): 351.
[88] Veeramachaneni K, Vladislavleva K, Burland M, Parcon J, O' Reilly Un-M. Evolutionary optimization of flavors. In: Proceedings of the 12th annual conference on Genetic and evolutionary computation, 2010; pp.1291–1298. doi: 10.1145/1830483.1830713.
[89] Queiroz L P, Rebello C M, Costa E A, Santana V V, Rodrigues B C L, Rodrigues A E. Generating flavor molecules using scientific machine learning. ACS Omega, 2023; 8(12): 10875–10887.
[90] Schreurs M, Piampongsant S, Roncoroni M, Cool L, Herrera-Malaver B, Vanderaa Christophe et al. Predicting and improving complex beer flavor through machine learning. Nature Communications, 2024; 15(1): 2368.
[91] Shen C, Cai G H, Tian J Q, Wu X N, Ding M Q, Wang B, et al. Characterization of lamb shashliks with different roasting methods by intelligent sensory technologies and GC–MS to simulate human muti-sensation: Based on multimodal deep learning. Food chemistry, 2024; 440: 138265.
[92] Singh R, Nisha R, Naik R, Upendar K, Nickhil C, Deka S C. Sensor fusion techniques in deep learning for multimodal fruit and vegetable quality assessment: A comprehensive review. Journal of Food Measurement and Characterization, 2024; 18(9): 8088–8109.
[93] Feng L, Wu B H, Zhu S S, He Y, Zhang C. Application of visible/infrared spectroscopy and hyperspectral imaging with machine learning techniques for identifying food varieties and geographical origins. Frontiers in Nutrition, 2021; 8: 680357.
[94] Pu H B, Yu J X, Sun D-W, Wei Q Y, Wang Z. Feature construction methods for processing and analysing spectral images and their applications in food quality inspection. Trends in Food Science & Technology, 2023; 138: 726–737.