8(2): 27-37, 2025

Journal of Food and Bioprocess Engineering

Journal homepage: https://jfabe.ut.ac.ir

Review article

Advancements in AI-driven process optimization and quality control for edible oils in Industry 4.0

Fereshteh Ramezani [a], Tahere Razzaghi [b, *]

[a] *Quality Control and Quality Assurance Departments, Newsha Darian Agro-Industry Co., Tehran, Iran. Postal Code 1443955511*
[b] *Department of Food Science and Engineering, College of Agriculture and Natural Resources, University of Tehran, Karaj Campus, Karaj, Iran.*

A B S T R A C T

Traditional quality control approaches are often reactive, labor-intensive, and limited in scalability, responsiveness, and precision. In contrast, AI and ML are transforming the edible oil manufacturing industry. They enable process optimization, real-time monitoring, and advanced quality control in line with Industry 4.0. This study reviews recent research and applications of AI and ML in edible oil extraction processes and quality control. It focuses on optimizing extraction parameters and yield, minimizing impurities, and ensuring safety to enable sustainable, intelligent production. Advanced algorithms such as ANNs and ANFIS offer superior accuracy for optimizing extraction, predicting antioxidant content, and controlling processes compared to conventional methods. For quality control, AI has enabled rapid, nondestructive assessments of oil authenticity and oxidation. Technologies such as LF-NMR, combined with CNNs, are used. AIoT sensor-based systems integrate intelligent sensors, cloud platforms, and deep learning models such as LSTM, ANNs, and CNNs. These systems enable real-time monitoring of rancidity, as well as volatile gas emissions and color changes during storage. Other advanced AI-driven innovations include image-based defect detection using DMEOI datasets and infrared cameras for real-time inspection. Emerging techniques such as HSI with ML, BME688 gas sensors, voltammetric electronic tongues, and visual array sensors detect adulteration in pure and blended oils. FPGAs are also used for real-time detection of gutter oils. Despite these advances, widespread industrial adoption faces challenges. Key issues include data quality, privacy, cybersecurity, workforce skills, and integration with legacy systems. Addressing these data issues is a major concern for industry and academia.

Keywords: Artificial intelligence; Machine learning; Edible oils; Quality control; Industry 4.0.

Received 27 Oct 2025; Received in revised form 29 Jan 2026; Accepted 31 Jan 2026

1. Introduction

In complex and competitive marketplaces, quality control (QC) assures safety, regulatory compliance, and consumer trust. Effective QC reduces contamination, fraud, adulteration, and process variability risks (Liakos et al., 2025).

Adulteration, long-term storage, and improper processing often cause edible oils to oxidize, rancidify, and hydrolyze, affecting quality, nutrition, and safety. Oil oxidation, the main quality-deteriorating reaction, would severely impact its quality and safety during processing and storage. Oil oxidation degree is evaluated by

Abbreviations: AI: Artificial Intelligence; AIoT: Artificial Intelligence of Things; ANFIS: Adaptive Neuro-Fuzzy Inference System; ANN: Artificial Neural Network; CNNs: Convolutional Neural Networks; DMEOI: Dual-Modal Edible Oil Impurity; EVOO: Extra Virgin Olive Oil; FMCG: Fast-Moving Consumer Goods; FPGAs: Field-Programmable Gate Arrays; FT-NIR: Fourier-transform near-infrared; HSI: Hyperspectral Imaging; IoT: Internet of Things; LF-NMR: Low-Field- Nuclear Magnetic Resonance; LSTM: Long Short-Term Memory; ML: Machine Learning; 1D-CAE: One-Dimensional Convolutional Autoencoder; PCA-LDA: Principal Component Analysis-Linear Discriminant Analysis; PLS: Partial Least Squares; QC: Quality Control; RS: Raman Spectroscopy; RSM: Response Surface Methodology; SFE: Supercritical Fluid Extraction; SVM: Support Vector Machine; SVR: Support Vector Regression.

*Corresponding author.
E-mail address: *tahere.razzaghi@ut.ac.ir* (T. Razzaghi).
https://doi.org/10.22059/JFABE.2026.405093.1221

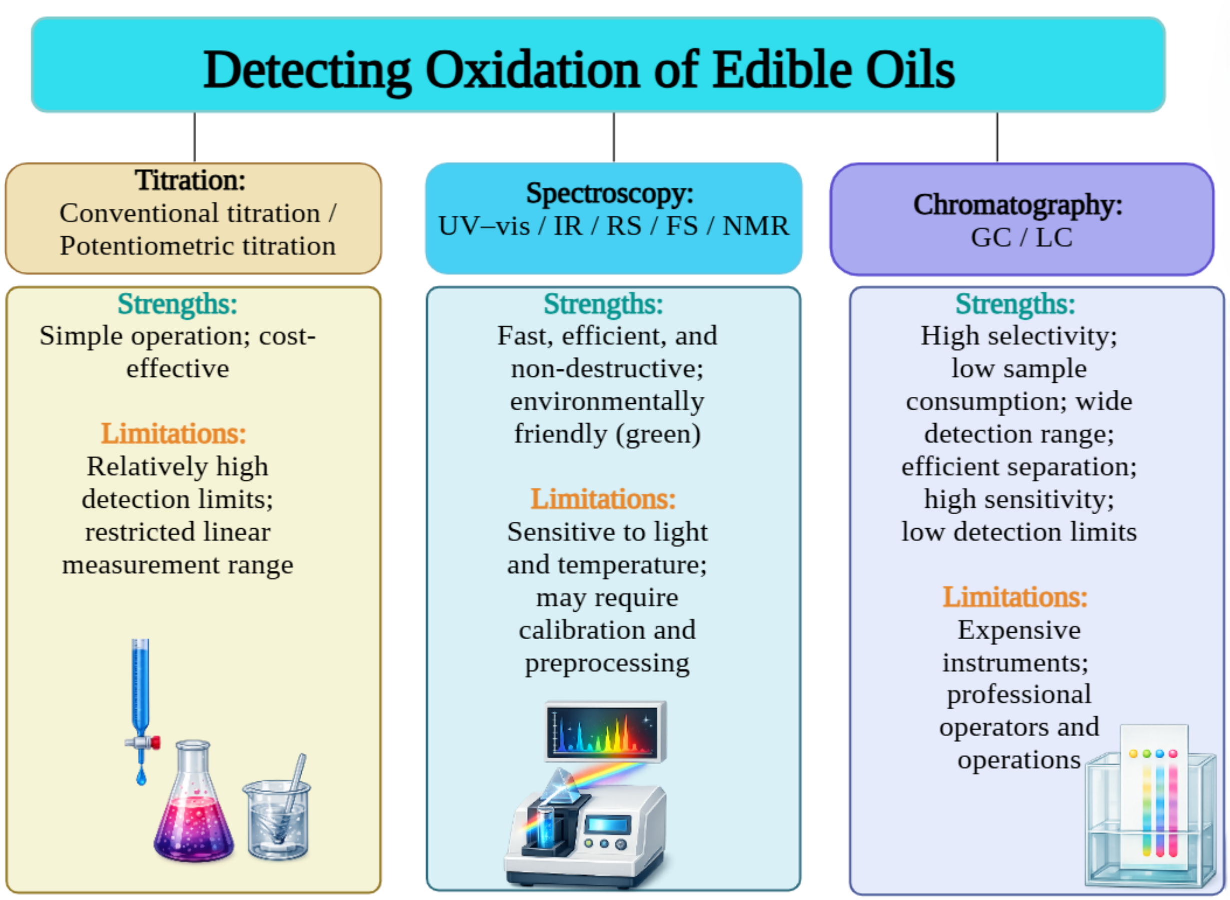

Fig. 1. Advantages and disadvantages of chromatography, titration, and spectroscopy methods for evaluating oil oxidation degree. The contents of this figure are compiled from (Lu et al., 2025). Abbreviations: UV–vis, UV–visible Spectroscopy; IR, Infrared Spectroscopy; RS, Raman Spectroscopy; FS, Fluorescence Spectroscopy; NMR, Nuclear Magnetic Resonance; GC, Gas Chromatography; LC, Liquid Chromatography.

multiple indexes, including peroxide value, acid value, total polar compounds, iodine value, 2-thiobarbituric acid, carbonyl value, and p-anisidine value.

Various analytical techniques have been suggested to evaluate the degree of oil oxidation by examining changes in the oil's sensory, chemical, and physical properties. Fig. 1 shows the advantages and disadvantages of each of these methods (Lu et al., 2025).

Traditional QC approaches, such as manual inspection, laboratory testing, and batch sampling, have been indispensable for decades but are often reactive, labor-intensive, and limited in scalability, responsiveness, and precision, particularly within the context of globalized supply chains and evolving consumer demands.

Currently, rapid detection methods for oil oxidation are emerging as a focal point in research and industry, as researchers, consumers, and producers seek prompt and practical access to oil quality information for effective safety alerts and control measures (Liakos et al., 2025).

In addition to quality monitoring, researchers have also been continuously striving to replace or complement traditional oil extraction methods, including solvent extraction, mechanical extraction, and the pre-press-solvent method. In these methods, the extraction efficiency depends on key factors such as the solvent used, the extraction technique, the characteristics of the plant component matrix, as well as process parameters including time, pressure, and temperature. Optimizing the extraction conditions and enhancing oil yield is desirable from both economic and environmental perspectives (Bakhshabadi et al., 2025).

Artificial intelligence (AI) algorithms significantly enhance industrial process control and monitoring. For example, statistical process control and statistical quality control methods require extensive prior knowledge of the process. AI-optimized process boundaries, on the other hand, provide valuable insights about the monitored process. AI has its requirements in industrial applications. A competitive advantage over conventional methods can be gained by using the appropriate AI algorithms to increase process knowledge, as well as prediction and forecasting capabilities. This evolution signifies a paradigm transformation in how quality assurance is conducted in industrial settings, particularly in the extraction of edible oils. Automatic quality prediction, full part traceability, process optimization, and preventive maintenance become possible with AI. These benefits directly influence productivity key performance indicators (KPIs), including overall equipment effectiveness and breakdown frequency (Bonada et al., 2020). Machine learning (ML) is highlighted as an emerging field in automated learning systems, characterized by its ability to develop algorithms that predict future behaviors by analyzing existing patterns in a dataset; particularly the supervised learning technique has shown to have significant effectiveness in industrial process optimization and in the agro-industrial process control (Márquez et

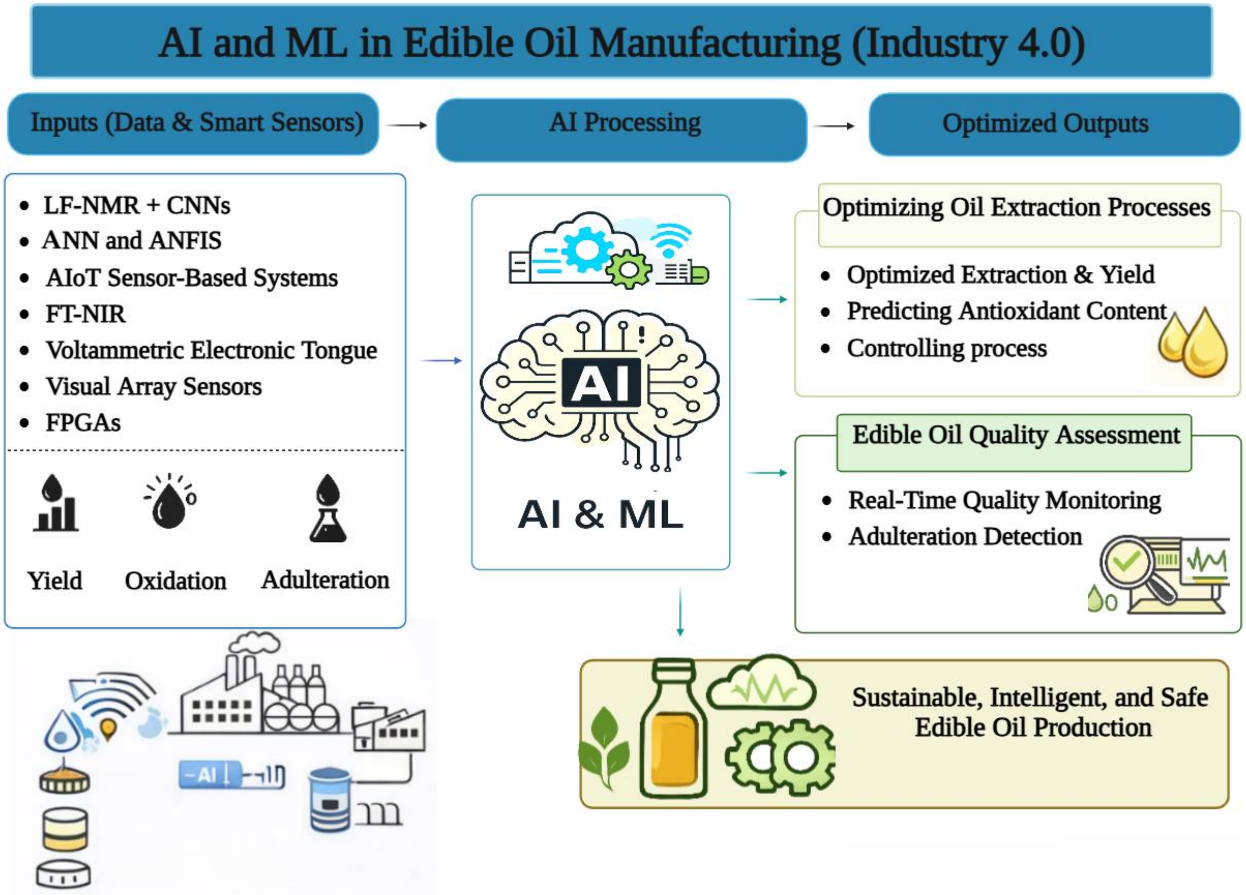

Fig. 2. Graphical overview of AI and ML applications in edible oil manufacturing under Industry 4.0. Abbreviations: AI, Artificial Intelligence; ML, Machine Learning; LF-NMR, Low-Field Nuclear Magnetic Resonance; CNNs, Convolutional Neural Networks; ANN, Artificial Neural Network; ANFIS, Adaptive Neuro-Fuzzy Inference System; AIoT, Artificial Intelligence of Things; FT-NIR, Fourier-Transform Near-Infrared; FPGAs, Field-Programmable Gate Arrays.

al., 2020). ML functions as a powerful instrument for providing very accurate predictions, classifying concepts, facilitating intelligent control, predicting maintenance needs, and detecting faults and anomalies in real time (Ramentol et al., 2021). The phrase "Industry 4.0" refers to the fourth industrial revolution, characterized by the application of digital technology to enhance automation, connectivity, and efficiency across diverse industrial processes.

AI facilitates optimal column designs, appropriate solvent selection, and extraction yield prediction. The combination of AI technology with extraction columns represents a significant step forward for industrial processes, bringing in a new era of intelligent systems that provide previously unheard-of enhancements in efficiency, quality, and environmental responsibility (Paneru et al., 2024). AI and ML are revolutionizing industries and businesses, where low-cost, big data can utilize computing power to optimize system performance. Technological improvements and the big data revolution can help provide the information needed for decision-making and reduce the time from identification to execution (Salem et al., 2022). These technological capabilities offer the industry a significant increase in efficiency, including energy savings, improved efficiency in human resources, improved product quality, and reduced environmental impact (Ramentol et al., 2021). This study reviews the practical implementation of AI and ML techniques in edible oil production, focusing on process optimization, quality control, and food safety improvements, and further discusses the existing challenges and future research directions in this field. Fig. 2 provides a graphical overview of the role of AI- and ML-based systems in optimizing edible oil production.

2. AI technologies for optimizing oil extraction processes

The broader Industry 4.0 paradigm encourages the use of smart sensors, devices, and machines to enable smart factories that continuously collect data about production (Rai et al., 2021). The integration of AI to enhance oil extraction processes signifies an important development in the edible oil industry (Fig. 3).

AI models can manage large volumes of data to enable real-time monitoring and management of extraction processes. Sensors track key parameters like pressure, temperature, and solvent flow rate continuously, while AI algorithms analyze this real-time data and autonomously adjust conditions to maintain efficiency and consistency. The algorithm learns from new data, identifying patterns and suggesting process refinements and efficiency enhancements in the extraction process. Additionally, ML methods facilitate predictive maintenance of equipment, minimizing downtime and ensuring continuous output in large-scale production (Alloun et al., 2024).

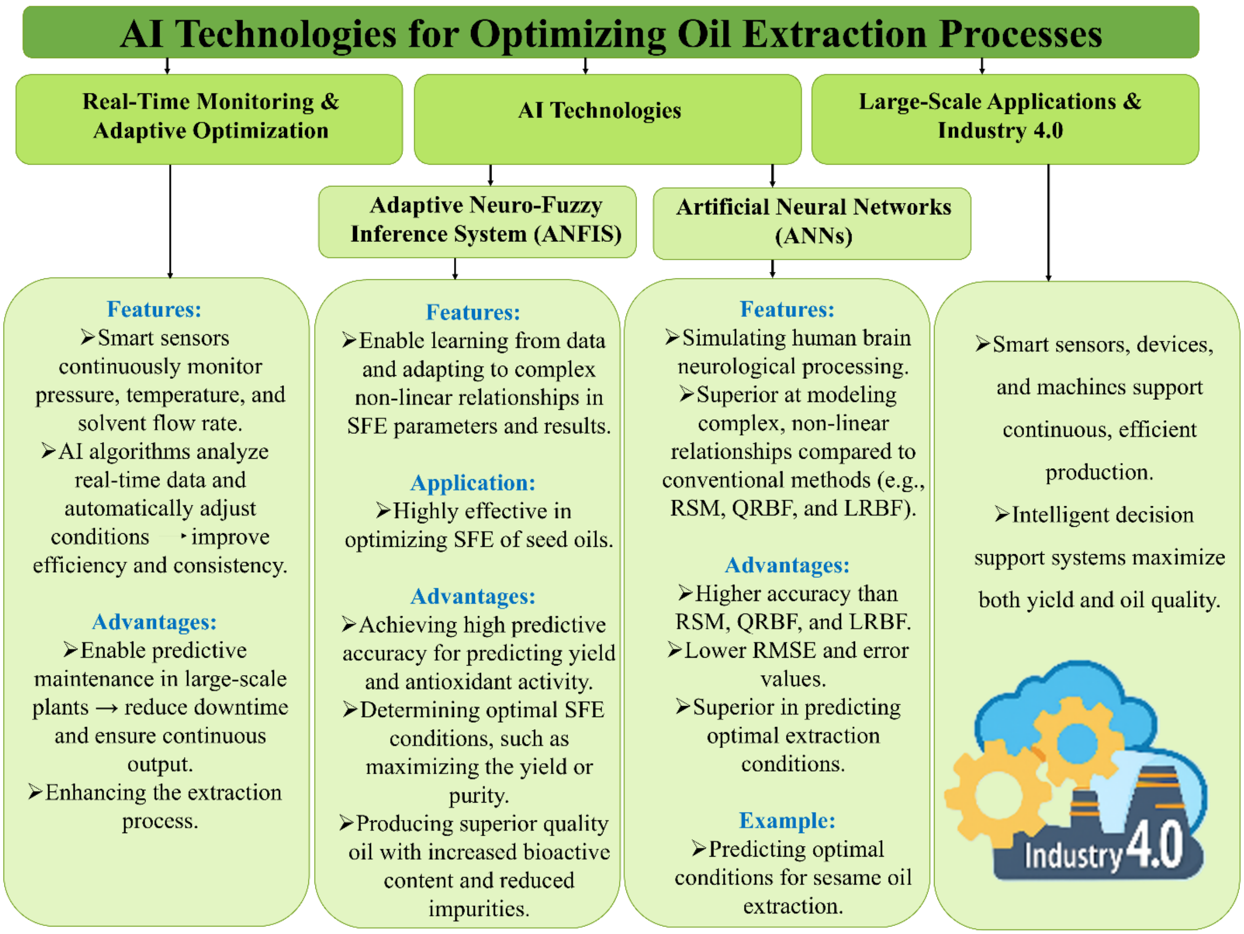

Fig. 3. Role of AI technologies in optimizing oil extraction processes. The contents of this figure are compiled from (Okunola et al., 2015; Osman et al., 2019; Pao-la-or et al., 2024; Alloun et al., 2024; Rai et al., 2021). Abbreviations: RSM, Response Surface Methodology; QRBF, Quadratic Radial Basis Functions; LRBF, Linear Radial Basis Functions; SFE, Supercritical Fluid Extraction.

This capability aligns with the Industry 4.0 paradigm (Rai et al., 2021). The following examples discuss how intelligent models are applied in oil extraction processes.

An artificial neural network (ANN) is a computational learning system that simulates human brain neurological processing and quantifies a non-linear relationship between connecting factors and actual responses using iterative training of data. ANN is superior for modeling data sets showing non-linear correlations and, thus, enabling both data fitting and prediction. This system has been applied in numerous fields, including biology, psychology, medicine, metrology, neurology, science, mathematics, and engineering disciplines (Okunola et al., 2015).

Several studies indicate that ANN optimizes oil extraction parameters better than conventional techniques such as Response Surface Methodology (RSM). For example, Okunola & Adepoju (2015) showed that ANN predicted ideal conditions for sesame (*Sesamum indicum*) oil extraction, yielding 85.70% oil compared to RSM (83.20%). Additionally, it seems that ANN was a more effective optimization tool than RSM based on its error value performance indicators (RMSE, AAD, R^2, R^2_{Adj}) (Okunola et al., 2015).

This conclusion was validated and further supported by an independent comparative study examining RSM, Quadratic Radial Basis Functions, Linear Radial Basis Functions, and ANN. These models were evaluated by monitoring their ability to predict the optimum conditions for sesame seed oil extraction. Model prediction efficiency was assessed using error value performance indicators (R^2, R^2_{Adj}, and RMSE). The ANN model consistently yielded more accurate predictions across all tested solvents and operating conditions, with a notably low RMSE (Osman et al., 2019). Therefore, ANN consistently outperforms conventional methods such as RSM in optimizing oil extraction parameters due to its non-linear modeling capabilities, higher yield, and low prediction error.

The Adaptive Neuro-Fuzzy Inference System (ANFIS) is another powerful AI tool for optimizing complex supercritical fluid extraction (SFE) parameters of seed oils. Although both RSM and ANFIS models are valuable tools, studies that further compared these optimization models showed ANFIS-Pomegranate Seed Oil achieves the lowest final free fatty acid content, and ANFIS potentially offers superior optimization capability.

According to Pao-la-or et al. (2024) ANFIS models achieve $R^2 > 0.99$ for predicting yield and antioxidant activity/purity and determine optimal SFE conditions for maximizing the yield (35 MPa, 45°C) or purity (30 MPa, 40°C).

By learning from data and adapting to the intricate non-linear relationships between the desired results and SFE parameters, ANFIS enhances extraction efficiency, yielding higher quality oil with increased bioactive content and reduced impurities (Pao-la-or et al., 2024).

The intelligent decision support systems allow edible oil producers to optimize various aspects of their operations at once, addressing the complex challenges of maximizing both yield and quality in extraction processes.

Although future studies should carefully investigate how the choice of optimization strategies and extraction methods affects fatty acid profiles, quality, potential health benefits, the antioxidant content, and also their cost-effectiveness and scalability for successful industrial implementation (Pao-la-or et al., 2024). A comparative summary of the main findings is presented in Table 1.

3. AI technologies for edible oil quality assessment

Assessing conventional oil quality mostly relies on chemical indicators such as peroxide value, acid value, and p-Anisidine value. This process requires expertise and involves complex procedures (Ku et al., 2023). In response to these limitations, researchers have developed AI-powered systems that can rapidly detect quality issues without disrupting production processes.

Assessing oil oxidation during production is crucial for monitoring purity and nutritional quality during processing and storage. The industry seeks a real-time, nondestructive, fast, robust, and cost-effective sensor for this purpose. A significant advancement in this field is the application of low-field nuclear magnetic resonance (LF-NMR) technology combined with convolutional neural networks (CNNs) for detecting oxidation. This AI-based LF-NMR real-time oxidation monitoring system achieves a 95% classification accuracy in distinguishing between non-oxidized, partially oxidized, and highly oxidized oils, while providing rapid results within several minutes using a relatively low-cost LF-NMR magnetometer and a fast-processing system, in contrast to conventional common chemical and spectral methods that takes days to weeks (Osheter et al., 2023).

A Sesame Oil Quality Assessment Service Platform, including an Intelligent Sesame Oil Evaluator (ISO Evaluator) and a Cloud Service Platform, was developed to quickly evaluate the quality of sesame oil. The ISO Intelligent Sesame Oil Evaluator employs Artificial Intelligence of Things (AIoT) sensors to monitor alterations in oil color and volatile gas emissions during storage, utilizing deep learning mechanisms, including Long Short-Term Memory (LSTM), CNNs, and ANN (Ku et al., 2023).

Similarly, researchers have developed a system that integrates a rancidity sensor with an Internet of Things (IoT) communication module, enabling real-time collection and prediction of rancidity values at a site. The sensor's measurements are periodically transferred to the cloud through the IoT communication module, the validity of the data set is verified, and systematic data management is performed using machine-learning-based data analysis approaches. The experimental results showed a high classification prediction accuracy of 91.3% and a short-term pattern prediction accuracy of 96.6% (weighted scaling), confirming its significant potential for raw material quality management (Hong et al., 2019).

Artificial vision has various applications for food quality control. This fast, affordable, consistent, and objective nondestructive technique has expanded in various industries. It satisfies ever-increasing production and quality requirements thanks to its speed and accuracy, thus helping develop completely automated processes. Over the last five years, image processing approaches have been designed for real-time control and automated inspection using different instrumentation (Guzmán et al., 2013).

Computer vision-based impurity detection methods, such as the dual-modal edible oil impurity (DMEOI) dataset, significantly reduce the workload while enhancing detection efficiency and accuracy. The DMEOI dataset has been developed and is publicly available for the detection of solid impurities in edible oils (Wang et al., 2024). Similarly, an automated defect detection system based on information extraction from olive images has been developed for olive oil production. An infrared camera is installed in the reception yard of the mill. The images of the olives are used for the detection of defects on individual olives (Marchal et al., 2021).

AI algorithms can handle complex, high-dimensional datasets, enabling manufacturers to gain a holistic view of the manufacturing process. Computer vision and ML algorithms have played a significant role in achieving these outcomes and optimizing operations (Verma, 2018).

Various statistical and machine-learning-based classification methods have been developed to detect edible oil adulteration, particularly in complex mixtures with multiple adulterants. For example, Dou et al. (2024) proposed a one-class classification and outlier detection approach based on Monte Carlo sampling. Their method can identify both single and multiple adulterations in peanut and sesame oils. By using this machine-learning-based technique, a novel solution emerges for inspecting potential adulteration in practice (Dou et al., 2024).

Hyperspectral imaging (HSI) uses the unique spectral signatures of different oil types and their mixtures to rapidly and non-destructively distinguish pure oils from adulterated ones. With further refinements and validation, HSI can become a key tool against food fraud and adulteration. Accordingly, Aqeel et al. (2024) identified substandard oil adulteration at different stages using the non-destructive HSI Specim Fx 10 system equipped with a Savitzky-Golay filter to remove noise and smooth spectral signatures. The proposed HSI approach, based on the HSI Specim FX10 system (400–1000 nm), provides spatial and spectral data for both pure and contaminated edible oils, enabling non-destructive, instantaneous fraud detection and classification. By capturing extensive spectral information for every pixel across a wide wavelength range, HSI allows a detailed understanding of the physical and chemical composition of edible oils and effectively distinguishes genuine from adulterated samples. The model achieved a validation accuracy of 100%, outperforming several state-of-the-art models, and demonstrates HSI as a precise and efficient tool for quality assurance and control in the food processing sector (Aqeel et al., 2024).

Based on recent imaging-based advancements, Hwang et al. (2024) applied HSI combined with machine learning to classify eight edible vegetable oils, compared with the chemical method based on fatty acid composition, and also used the method to ascertain adulteration levels in binary oil blends. The hyperspectral absorbance spectra showed that palm oil exhibited distinct spectral features due to its high degree of saturation, while flaxseed and olive oils showed dominant hyperspectral intensities at distinct wavelengths (1170/1671 nm and 1212/1415 nm, respectively). Using Linear Discriminant Analysis (LDA), the researchers showed that a significant portion of the total variability between the oils could be explained by two linear discriminants: fatty acid composition and hyperspectral image data, which explained 96% and 98.9% of the differences, respectively. ML models achieved high classification performance, surpassing 98.9% on the hyperspectral images of vegetable oils, comparable to the chemical method based on fatty acid composition for detecting edible

Table 1. Summary of AI-based approaches for optimizing oil extraction processes.

Modeling techniques	Process focus	Performance indicators	Key findings	Strengths	Limitations	References
ANN, RSM	Sesame oil extraction	RMSE, AAD, R^2, R^2_{Adj}	ANN achieved higher yield (85.7%) than RSM (83.2%) with lower error values.	Produced high-quality, highly unsaturated edible oil with potential industrial applications	- Real-scale and multi-factor validation not performed. - Comparison restricted to RSM model.	(Okunola & Adepoju, 2015)
ANN, QRBF, LRBF, RSM	Multi-solvent sesame oil extraction under various conditions	R^2, R^2_{Adj}, RMSE	The trained ANN showed superior prediction accuracy compared to other models.	- Comprehensive multi-model comparison; - Effective for nonlinear system modeling	Comparison restricted to controlled lab conditions	(Osman et al., 2019)
ANFIS	SFE of seed oils	$R^2 > 0.99$	Accurately identified optimal SFE parameters for either yield (35 MPa, 45°C) or antioxidant purity (30 MPa, 40°C).	- Learns complex nonlinear relationships. - Enhances extraction efficiency and bioactive content while reducing impurities.	Specific to SFE process	(Pao-la-or et al., 2024)

vegetable oils. For ascertaining adulteration levels in binary oil blends, the Random Forest model was the most effective ($R^2 > 0.992$; RMSE < 2.75). This study showed that combining hyperspectral imaging with machine learning is a highly accurate method for identifying and controlling potential quality issues in vegetable oils (Hwang et al., 2024).

Economically motivated adulteration and cross-contamination are the primary reasons involved in the inferior quality of premium edible oils, such as extra virgin olive oil (EVOO), in the food and beverage and fast-moving consumer goods (FMCG) industries. Current quality assurance methods in FMCG, such as spectroscopy and chromatography techniques, have their limitations, including the requirement of intrusive sample extraction and ex-situ laboratory analysis with expensive, bulky instrumentation that is neither scalable to match production throughput nor integrable inline. These methods do not meet industrial requirements for in situ, non-intrusive, high-throughput inspection, leading to food loss and package waste from undesired batch rejects (Mohan et al., 2023).

Recent advances have demonstrated that spectroscopy-based techniques can achieve much higher precision when combined with machine learning and deep learning models. For instance, Lim et al. (2025) developed an explainable AI-based models combining ^{1}H NMR spectroscopy with multivariate and deep learning algorithms, including principal component analysis-linear discriminant analysis (PCA-LDA), support vector machine (SVM), and 1D-CNN as classification models to differentiate among various edible oil types, and PLSR, support vector regression (SVR), and 1D-CNN regression as regression models to quantify sesame oil content in adulterated oil mixtures. 100% accuracy was achieved by all three classification models (PCA-LDA, SVM, and 1D-CNN), with 1D-CNN also capturing both lignan and fatty acid signals. 1D-CNN regression model (RMSEP = 1.03, $R^2 = 0.999$) also demonstrated superior performance than PLSR and SVR models (RMSEP = 1.94 and 1.40, $R^2 = 0.998$, respectively). Moreover, the CNN model accurately predicted sesame oil content for previously unused oil types within a 2% error margin. These findings highlight the potential of AI-enhanced NMR spectroscopy for detecting adulteration of edible oils (Lim et al., 2025).

Similarly, a recent study introduced a novel approach for detecting butylated hydroxytoluene, a common antioxidant in edible oils, by combining a one-dimensional convolutional autoencoder (1D-CAE) model with Fourier-transform near-infrared (FT-NIR) spectroscopy. The 1D-CAE compressed the spectral data into condensed features, which were then analyzed using support vector machine and partial least squares (PLS) regression models to quantify butylated hydroxytoluene levels. This method demonstrated remarkable repeatability and high prediction accuracy ($R^2 = 0.9953$; RMSE = 1.2035; RPD = 15.1664) and can be utilized to construct robust detection models (Deng et al., 2025).

A novel hybrid approach has been proposed for detecting adulteration in edible oils by combining the value of refractive index with electronic sensors trained with algorithms based on machine learning, like k-nearest neighbors (k-NN), Random Forest, CATBOOST, and XGBOOST, for accurately revealing the adulteration. In the first method, changes in the refractive indices of pure vegetable oil, which are decreased by vegetable oils and increased by animal fats, are measured using advanced attenuated total reflectance mid-infrared (ATR-MIR) Spectroscopy. These spectral data are considered for detecting adulteration. In the second method, adulterated and pure oil samples are heated at 40–50 °C for 4.66 min by a Fumigation technique that integrates real-time hardware. The generated volatiles were transmitted to the MEMS Gas Sensor-MISC-2714 and the Multichannel Gas Sensor. This approach utilizes the conductance value as a feature for the classifier to determine if the sample is severely, moderately, or minimally contaminated. Finally, among the evaluated the ML algorithms, the Random Forest Classifier and the XGBoost algorithm performed well, achieving 100% accuracy (Agrawa et al., 2024).

While NMR-, ATR-MIR-, and FT-NIR-based AI systems demonstrate exceptional analytical precision (Lim et al., 2025), their implementation in industrial environments remains limited due to high instrumentation costs and complex data-processing requirements (Mohan et al., 2023). Consequently, complementary low-cost, rapid sensing methods have been developed to enable real-time in-line quality assurance.

For industrial quality assurance, a low-cost capacitance-based electrical approach has been developed for screening EVOO-filled containers non-invasively for adulteration without any sample extraction. This technology, which is suitable for high-throughput (>60 samples/min) screening, detects the differences in the dielectric characteristics of blended oils. The sensor system demonstrates a quick response time (100 ms) and low detection limits for different adulterants, including olive oil (32.8%), canola oil (19.4%), soy oil (10.3%), and castor oil (1.7%). In contrast, conventional industrial methods such as spectroscopy and chromatography typically show response times ranging from seconds to hours (5 s - 1.5 h), detection limits below 1%, and require sample extraction, offline analysis, and higher costs (Mohan et al., 2023).

Beyond capacitance-based methods, several innovative sensor-driven and AI-enhanced approaches have also been proposed for real-time adulteration detection. In a recent study, Parlak et al. (2025) employed a portable BME688 gas sensor combined with machine learning algorithms to detect adulteration in extra virgin olive oil. Sunflower oil was mixed into extra virgin olive oil at varying concentrations, and the gases released from these mixtures were analyzed using the high-sensitivity sensor. Both classification and regression models achieved near-perfect accuracy in detecting adulteration levels. However, the study’s limitations include controlled experimental conditions, testing only sunflower seed oil as an adulterant, and using a single type of electronic nose sensor, suggesting that future work should validate the system under real-world conditions and with multiple sensor combinations. Overall, this low-cost, portable sensing system demonstrates strong potential as a rapid and scalable method for olive oil authentication (Parlak et al., 2025). Other sensor-driven AI methods have also been explored for real-time adulteration detection in various edible oils.

In contrast to human tongues, voltammetric electronic tongues can serve quantitative functions when their signals undergo multivariate calibration techniques, such as PLS or ANN. In food analysis, predictions can be made regarding numerous characteristics that indicate the quality of food products. In this regard, Apetrei (2012) assessed the quality of six extra virgin olive oils based on their bitterness levels and employed a multivariate calibration model to predict the associated bitterness indices (Pérez-Ràfols et al., 2019; Apetrei 2012).

Also, according to Men et al. (2013) voltammetric electronic tongue systems enhanced with artificial fish swarm algorithms can optimize the cluster centers' value and the initial value, which is important for the detection and classification of edible oil using the artificial fish swarm cluster algorithm in a voltammetric electronic tongue system (Men et al., 2013).

Similarly, ML algorithms are deployed on Field-Programmable Gate Arrays (FPGAs) for real-time gutter oil detection under multiple feature dimensions. It is deployed on an FPGA to be low-power and portable for actual use, which addresses excessive limitations of laboratory tests (Jiang et al., 2021). Additionally, a visual array sensor for detecting sesame oil adulteration was developed, achieving a 100% classification accuracy (Liu et al., 2022). A comparative summary of AI-based approaches for edible oil quality assessment is presented in Table 2.

4. Challenges and future research directions

The challenges associated with AI adoption in the edible oil industry can generally be categorized into two key areas: data-related challenges and implementation barriers. These domains collectively shape the extent to which AI technologies can be effectively deployed and sustained within industrial environments.

4.1. Data management challenges

Although AI adoption has gained considerable attention across various industries, its actual level of adoption remains limited, with data-related challenges considering as a one of the major barriers to effective AI implementation (Li et al., 2024). In this context, data-related challenges can be discussed in two major areas: the need for high-quality, reliable data and concerns about data privacy and security.

4.1.1. The need for quality data

AI systems heavily rely on large volumes of high-quality data to operate effectively (Rakholia et al., 2024). In fact, in industrial systems, data is the basis on which AI algorithms operate. Implementing AI requires collecting relevant data, storing it appropriately, ensuring data quality through cleaning, and analyzing and visualizing data for the intended purposes (Li et al., 2024).

Whereas some manufacturing companies may lack the necessary data infrastructure or have incomplete, outdated, or low-quality data. Without adequate data, AI systems may not achieve the desired level of accuracy and efficiency (Rakholia et al., 2024).

4.1.2. Data privacy and security concerns

Ethical concerns related to data privacy, algorithmic bias, and the potential for cyber-attacks are challenges that must be carefully addressed to protect intellectual property and sensitive information (Agrawal et al., 2025; Rakholia et al., 2024). Addressing these data issues is a significant concern for both industry and academia (Li et al., 2024). Beyond data-related limitations, industries also encounter several implementation barriers that impede the large-scale integration of AI systems into existing production frameworks, as discussed in section 4.2.

4.2. Implementation Barriers

Despite rapid advancements and growing AI adoption in the food industry, food manufacturers face challenges in developing, deploying, and maintaining AI systems. These challenges include difficulty integrating AI with legacy systems, risk aversion toward new technologies, lack of clear implementation strategies, shortage of skilled professionals, high costs, the need for standardized AI governance frameworks, and limited investment (Agrawal et al., 2025; Pennells et al., 2025; Rakholia et al., 2024; Verma, 2018).

These challenges hinder the widespread adoption of AI in food manufacturing, and can create a cultural barrier to AI adoption within the organization (Agrawal et al., 2025; Rakholia et al., 2024). Facilitating large-scale screening of edible oil samples and enabling faster and more accurate analysis necessitate the development of AI algorithms, improvements in data analysis techniques, a skilled workforce, and a cultural shift that promotes data-driven decision-making and promotes harmonious collaboration between human expertise and machine capabilities (Rakholia et al., 2024; Srivastava et al., 2024).

Table 2. Summary of AI-based approaches for edible oil quality assessment.

	Modeling Techniques / AI Tools	**Process Focus**	**Performance Indicators**	**Key Findings**	**Strengths**	**Limitations**	**References**
Oxidation and Authenticity	**1H LF-NMR relaxation sensor + CNN**	Oxidation status monitoring	Classification accuracy: 95%	Classifies oils into non-oxidized, partially oxidized, and highly oxidized using T2 relaxation curves	- Semi-autonomous, real-time, non-destructive, rapid, and low-cost method, - Worked with large-scale dataset	Requires LF-NMR sensor and a large training dataset	(Osheter et al., 2023)
	AIoT sensors + LSTM + CNN + ANN	Sesame oil quality assessment (volatile gases and color changes)	LDA: 95.13; CNN+ANN MAPE: 8.18%; LSTM MAPE: 0.44%	Rapid evaluation of oil quality, predict future changes, and monitor storage conditions	- Real-time, cloud-based, scalable, digitized quality monitoring - Predictive capability	-	(Ku et al., 2023)
	Rancidity sensor + IoT + ML	Real-time rancidity prediction	Classification accuracy: 91.3%; Short-term pattern prediction: 96.6%	IoT sensor + ML enables real-time rancidity data collection and analysis.	Excellent potential for raw material quality management	Currently applied to road pavement asphalt	(Hong et al., 2019)
Impurities and Adulteration	**Computer vision-based detection using DMEOI dataset**	Detection of tiny solid impurities in edible oils	Dataset of 14,520 images; Object detection algorithms applied for validation	DMEOI dataset allows both single- and dual-modal detection of liquid/solid impurities in oils.	Reduces manual workload; Improves detection efficiency and accuracy; Publicly available dataset	Real-time industrial deployment not reported	(Wang et al., 2024)
	Automatic defect detection system using infrared imaging	Detection of defective olives in oil mill reception area	High detection rate of individual fruit defects	Infrared imaging + active contour + decision tree effectively identifies defective olives	Suitable for the virgin olive oil industry	Limited to olive fruit	(Marchal et al., 2021)
	One-class classification + outlier detection with Monte Carlo sampling	Detection of single and multiple adulterations in in peanut and sesame oils	Validation on three datasets	Effectively identifies adulterated samples	Effective practical multiple- adulterants detection	-	(Dou et al., 2024)
	HSI + ML	Adulteration detection and classification	Validation accuracy: 100% Linear discriminant analysis (LDA) explained 96-98.9% of variability; Random Forest for adulteration levels ($R^2 > 0.992$, RMSE < 2.75)	Distinguish pure vs adulterated oils, classify vegetable oils, predict binary blend adulteration	Non-destructive; Rapid; Applicable for complex oil mixtures	-	(Aqeel et al., 2024) (Hwang et al., 2024)
	^{1}H NMR + PCA-LDA, SVM, 1D-CNN (classification); PLSR, SVR, 1D-CNN (regression)	Classification of edible oil types & quantification of sesame oil in adulterated mixtures	Classification Accuracy: 100% (all models); 1D-CNN Regression: RMSEP: 1.03, R^2: 0.999; Prediction error for new oils: ±2%	1D-CNN accurately detects and quantifies sesame oil, capturing both lignan and fatty acid signals	High precision; robust to new/unseen oil types; integrates broad spectral features	-	(Lim et al., 2025)
	BME688 gas sensor + ML	Olive oil adulteration detection (sunflower oil in EVOO)	Near-perfect accuracy	Detect sunflower oil in EVOO blends using gas emission data	Portable, low-cost, rapid, and scalable method	Controlled experimental conditions; Single adulterant (sunflower oil); One sensor type only	(Parlak et al., 2025)

Continued.

Continued.

	FT-NIR + 1D Convolutional Autoencoder (1D-CAE) + SVM/PLSR	Detection and quantification of BHT in edible oils	R^2: 0.9953; RMSE: 1.2035; RPD: 15.1664	1D-CAE compressed spectral data into condensed features enabling highly accurate and repeatable BHT detection	High prediction accuracy and repeatability	-	(Deng et al., 2025)
	ATR-MIR / Fumigation + MEMS Gas Sensors + ML (KNN, Random Forest, CATBOOST, XGBOOST)	Hybrid detection of adulteration in edible oils using spectral and volatile features	Random Forest & XGBOOST achieved 100% accuracy	Combined refractive index variation and gas-sensor features enabled accurate classification of adulteration levels (high, moderate, low)	Dual-modality sensing (spectroscopic + gas sensor); Real-time and hardware-integrated detection; Extremely high model accuracy (100%)	Requires controlled heating and calibration for reproducibility	(Agrawa et al., 2024)
	Capacitance-based electrical sensing system	Adulteration screening of EVOO in containers	Response time: 100 ms; Detection limits: olive 32.8%, canola 19.4%, soy 10.3%, castor 1.7%; Throughput > 60 samples/min	Fast, low-cost, non-invasive, and automated detection with low limits	Low-cost, fast, suitable for inline QA	Limited to proof-of-concept stage	(Mohan et al., 2023)
	Voltammetric electronic tongue + AFSA	Edible oil classification and detection	Optimization of cluster centers and initial values	Improved AFS algorithm effectively classifies edible oil samples	Global optimization for large-scale combinatorial problems	-	(Men et al., 2013)
Impurities and Adulteration	**Voltammetric electrodes based on polypyrrole (Ppy)**	Discrimination of the six extra virgin olive oils based on their degree of bitterness	Cyclic voltammograms: redox properties of electroactive compounds (mainly antioxidants) in olive oil emulsions	Successful discrimination using principal component analysis (PCA) and the partial least squares discriminant analysis (PLS-DA)	High cross-selectivity; Good correlation of electrochemical responses with physico-chemical and sensorial characteristics	-	(Apetrei, 2012)
	E-nose system based on 5-TGS sensors + E-tongue formed by 7-voltammetric electrodes	Argan oil adulterated with sunflower oil	Three pattern recognition methods: PCA, discriminant factor analysis (DFA), and SVMs	SVM (e-nose data): - 91.67% for comestible argan oil - 83.34% for cosmetic argan oil; E-tongue system: Perfect recognition rate using PCA, DFA, and SVM (up to 100% with SVM)	Accurate discrimination of pure, adulterated, and sunflower oils (10–70%)	Limited to sunflower oil as the adulterant	(Bougrini et al., 2014)
	Voltammetric e-tongue (based on modified carbon paste-based sensors)	Adulteration EVOOs with different percentages of sunflower oil, soybean oil and corn oil based on their tasting fingerprints	Kernel-based processing of square-wave voltammetric signals; Chemometric methods for discrimination and classification of oils according to botanical origin	Strong correlation between voltammetric signals and polyphenolic content (PLS); Feasible detection of olive oil adulteration below 10% using PLS and PLS-DA regression	Detection of adulteration below 10%; Accurate classification of multiple seed oils	Limited to seed oil adulterants	(Apetrei & Apetrei, 2014)
	FPGA + ML	Gutter oil detection	Classification accuracy: 97.18%; Processing time: 4.77 μs; Power consumption: 65.62 mW	k-nearest neighbors (k-NN) algorithm achieved highest classification accuracy; real-time detection feasible; low-power FPGA deployment	Real-time, low-power, portable, multi-feature detection; open-source dataset and code	-	(Jiang et al., 2021)
	Visual array sensor + ML (DD-SIMCA & PLSR)	Sesame oil adulteration detection	Classification accuracy: 100%	DD-SIMCA model effectively classifies pure vs. adulterated sesame oil; PLSR quantifies adulteration	Rapid, flavor-compound-based, non-destructive authentication	Limited to specific flavor compounds	(Liu et al., 2022)

4. Conclusion

This review emphasizes the transformative impact of AI and ML on edible oil production through process optimization, real-time monitoring, and intelligent quality control. The industrial adoption of AI technologies offers substantial improvements in efficiency, including energy savings, optimized use of human resources, enhanced product quality, and a reduced environmental footprint. AI algorithms can analyze both structured and unstructured data from many sources, including sensors, production logs, and maintenance records, facilitating real-time decision-making and predictive analytics. Despite these advancements, challenges remain in data quality, privacy, and cybersecurity, as well as the need for skilled professionals and seamless integration with existing systems. Overcoming these barriers is essential to fully realize the benefits of smart manufacturing and ensure the secure and ethical deployment of AI technologies. Technological advances and big data can significantly enhance decision-making efficiency, reducing the time between identification and implementation. Looking toward the future, the development of more robust AI algorithms and advanced data analysis techniques will enable the standardization and automation of biomarker-based authentication procedures, facilitating large-scale screening of edible oil samples and enabling faster, more accurate assessments. Overall, integrating AI and ML into edible oil production not only improves existing quality control practices but also aligns with Industry 4.0 principles, supporting safer, more sustainable, and data-driven food systems.

Acknowledgements

The support of the University of Tehran is gratefully acknowledged. The authors also sincerely thank Dr. Maryam Salami for her continuous support throughout this work.

Conflict of interest

The authors declare that there is no conflict of interest.

References

Agrawal, K., Goktas, P., Holtkemper, M., Beecks, C., & Kumar, N. (2025). AI-driven Transformation in Food Manufacturing: A Pathway to Sustainable Efficiency and Quality Assurance. *Frontiers in nutrition*, *12*, 1553942. https://doi.org/10.3389/fnut.2025.1553942

Agrawal, U., Bawane, N., Alsubaie, N., Alqahtani, M., Abbas, M., & Othman, S. (2024). Design & Development of Adulteration Detection System by Fumigation Method & Machine Learning Techniques. *Scientific Reports*, 14(1). DOI:10.1038/s41598-024-64025-4

Alloun, W., & Calvio, C. (2024). Bio-Driven Sustainable Extraction and AI-Optimized Recovery of Functional Compounds from Plant Waste: A Comprehensive Review. *Fermentation*, 10(3), 126. DOI:10.3390/fermentation10030126

Apetrei, C. (2012). Novel Method Based on Polypyrrole-Modified Sensors and Emulsions for the Evaluation of Bitterness in Extra Virgin Olive Oils. *Food Research International*, 48, 673-680. DOI:10.1016/J.FOODRES.2012.06.010

Apetrei, I.M., & Apetrei, C. (2014). Detection of Virgin Olive Oil Adulteration Using a Voltammetric E-Tongue. *Computers and Electronics in Agriculture*, 108, 148-154. https://doi.org/10.1016/j.compag.2014.08.002

Aqeel, M., Sohaib, A., Iqbal, M., Rehman, H.U., & Rustam, F. (2024). Hyperspectral Identification of Oil Adulteration Using Machine Learning Techniques. *Current research in food science*, *8*, 100773. DOI: 10.1016/j.crfs.2024.100773

Bakhshabadi, H., Ganje, M., Gharekhani, M., Mohammadi-Moghaddam, T., Aulestia, C., & Morshedi, A. (2025). A Review of New Methods for Extracting Oil from Plants to Enhance the Efficiency and Physicochemical Properties of the Extracted Oils. *Processes*, 13(4), 1124. https://doi.org/10.3390/pr13041124

Bonada, F., Echeverria, L., Domingo, X., & Anzaldi, G. (2020). AI for improving the overall equipment efficiency in manufacturing industry. In L. Romeral Martínez, R. A. Osornio Rios, & M. Delgado Prieto (Eds.), *New trends in the use of artificial intelligence for the industry 4.0* (Chap. 5). IntechOpen. DOI:10.5772/intechopen.89967

Bougrini, M., Tahri, K., Haddi, Z., Saidi, T., & EL Bari, N., & Bouchikhi, B. (2014). Detection of Adulteration in Argan Oil by Using an Electronic Nose and a Voltammetric Electronic Tongue. *Journal of Sensors*. 2014(4). DOI:10.1155/2014/245831

Deng, J., Chen, Z., Jiang, H., & Chen, Q. (2024). High-Precision Detection of Dibutyl Hydroxytoluene in Edible Oil via Convolutional Autoencoder Compressed Fourier-Transform Near-Infrared Spectroscopy. *Food Control*, 167. 110808. DOI:10.1016/j.foodcont.2024.110808

Dou, X., Tu, F., Yu, L., Yang, Y., Ma, F., Wang, X., Du, W., Zhang, L., Jiang, X., & Li, P. (2024). Adulteration Detection of Edible Oil by One-Class Classification and Outlier Detection. *Food Frontiers*, 5(4), 1806-1818. DOI: 10.1002/fft2.395

Guzmán, E., Baeten, V., Pierna, J., & García-Mesa, J.A. (2013). Using a Visible Vision System for On-Line Determination of Quality Parameters of Olive Fruits. *Food and Nutrition Sciences*, 04 (7A), 90-98. DOI:10.4236/FNS.2013.47A011

Hong, S., Chang, K., Lee, J., & Kim, B. (2019). Rancidity Analysis Management System Based on Machine Learning Using IoT Rancidity Sensors. *Sensors and Materials*, 31(11), 3871–3883. DOI:10.18494/sam.2019.2590

Hwang, J., Choi, K.O., Jeong, S., & Lee, S. (2024). Machine Learning Identification of Edible Vegetable Oils from Fatty Acid Compositions and Hyperspectral Images. *Current research in food science*, *8*, 100742. DOI: 10.1016/j.crfs.2024.100742

Jiang, W., Ma, Y., & Chen, R. (2021). Gutter Oil Detection for Food Safety Based on Multi-Feature Machine Learning and implementation on FPGA with Approximate Multipliers. *PeerJ Computer Science*, 7, e774. DOI:10.7717/peerj-cs.774

Ku, H., Lung, C., & Chi, C. (2023). Design of an Artificial Intelligence of Things-Based Sesame Oil Evaluator for Quality Assessment Using Gas Sensors and Deep Learning Mechanisms. *Foods*, 12(21), 4024. DOI:10.3390/foods12214024

Li, X., Cheng, Y., Møller, C., & Lee, J. (2025). Data Issues in Industrial AI System: A Meta-Review and Research Strategy. *Computers in Industry*, 173, 104361. https://doi.org/10.1016/j.compind.2025.104361

Liakos, K.G., Athanasiadis, V., Bozinou, E., & Lalas, S.I. (2025). Machine Learning for Quality Control in the Food Industry: A Review. *Foods*, 14(19), 3424. https://doi.org/10.3390/foods14193424

Lim, H., Cho, H., Kim, J.Y., & Shin, Y.J., Chun, H.S., & Kim, B.H., & Ahn, S. (2025). Classification And Quantification of Sesame Oil in Edible Oils and Adulterated Mixtures Using 1H NMR Spectroscopy Combined with Multivariate, Machine Learning, And Deep Learning Models. *Food Chemistry*, 493(4), 146008. DOI:10.1016/j.foodchem.2025.146008

Liu, R., Chen, H., Bai, X., Huang, Y., Li, H., Long, W., Lan, W., She, Y., & Fu, H. (2022). Visual Classification for Sesame Oil Adulteration Detection and Quantification of Compounds Used as Adulterants Using Flavor Compounds Targeted Array Sensor in Combination with DD-SIMCA and PLS. *Sensors and Actuators B: Chemical*. 357, 131335. DOI:10.1016/j.snb.2021.131335

Lu, Y., Xiong, R., Tang, Y., Yu, N., Nie, X., Zhang, L., & Meng, X. (2025). An Overview of the Detection Methods to the Edible Oil Oxidation Degree: Recent Progress, Challenges, and Perspectives. *Food Chemistry*, 463(4), 141443. https://doi.org/10.1016/j.foodchem.2024.141443

Marchal, P.C., Martínez, S.S., Ortega, J.G., & García, J.G. (2021). Automatic System for the Detection of Defects on Olive Fruits in an Oil Mill. *Applied Sciences*, 11(17), 8167. DOI:10.3390/APP11178167

Márquez, A.J., & Maza, G.B. (2020). Machine Learning Application in Process on Extra Virgin Olive Oil Elaboration Disk Stack Vertical Centrifuge Modeling. *International Journal of Computer Applications*, 176(37), 30-35. DOI:10.5120/ijca2020920511

Men, H., Zhang, C., Zhang, P., & Gao, H. (2013). Application of Electronic Tongue in Edible Oil Detection with Cluster Algorithm based on Artificial Fish Swarm Improvement. *Advance Journal of Food Science and Technology*, 5(4), 469-473. DOI:10.19026/AJFST.5.3293

Mohan, H.K.S.V., Aung, P.P., Ng, C.F., Wong, Z.Z., & Malcolm, AA. (2023). Rapid Non-Invasive Capacitive Assessment of Extra Virgin Olive Oil Authenticity. *Electronics*, 12(2), 359. DOI:10.3390/electronics12020359

Okunola, A.A., & Adepoju, T.F. (2015). Modeling and Optimization of Extraction of Oil from Sesamum Indicum Seeds: A Case Study of Response Surface Methodology vs. Artificial Neural Network. *International Journal of Chemistry and Materials Research*, 3(2), 41–52. DOI:10.18488/JOURNAL.64/2015.3.2/64.2.41.52

Osheter, T., Campisi Pinto, S., Randieri, C., Perrotta, A., Linder, C., & Weisman, Z. (2023). Semi-Autonomic AI LF-NMR Sensor for Industrial Prediction of Edible Oil Oxidation Status. *Sensors (Basel, Switzerland)*, 23(4), 2125. DOI:10.3390/s23042125

Osman, H., Shigidi, I., & Arabi, A. (2019). Multiple Modeling Techniques for Assessing Sesame Oil Extraction under Various Operating Conditions and Solvents. *Foods*, 8(4), 142. DOI:10.3390/foods8040142

Paneru, B., Paneru, B., & Chalise, A. (2024). Development of an Extraction Column using Machine Learning for the Prediction of Feed and Solvent to Obtain the Desired Extraction. *Journal of Artificial Intelligence and Capsule Networks*, 6(1), 27-44. DOI:10.36548/jaicn.2024.1.003

Pao-la-or, P., Marungsri, B., Chirinang, P., Posridee, K., Oonsivilai, R., & Oonsivilai, A. (2024). Boosting *Purnica granatum* L. Seed Oil Yield: An Adaptive Neuro-Fuzzy Interference System Fuels SC-CO2 Extraction Breakthrough. *Foods*, 13(1), 161. DOI:10.3390/foods13010161

Parlak, I.H., Milli, M., & Söylemez-Milli, N. (2025). Machine Learning–Based Detection of Olive Oil Adulteration Using BME688 Gas Sensor Matrix. *Food Analytical Methods*, 18, 1454-1464. DOI:10.1007/s12161-025-02803-0

Pennells, J., Watkins, P., Bowler, A.L., Watson, N.J., & Knoerzer, K. (2025). Mapping the AI Landscape in Food Science and Engineering: A Bibliometric Analysis Enhanced with Interactive Digital Tools and Company Case Studies. *Food Engineering Reviews*, 17, 465-489. https://doi.org/10.1007/s12393-025-09413-w

Pérez-Ràfols, C., Serrano, N., Ariño, C., Esteban, M., & Díaz-Cruz, J. M. (2019). Voltammetric Electronic Tongues in Food Analysis. *Sensors*, 19(19), 4261. https://doi.org/10.3390/s19194261

Rai, R., Tiwari, M.K., Ivanov, D., & Dolgui, A. (2021). Machine Learning in Manufacturing and Industry 4.0 Applications. *International Journal of Production Research*, 59(16), 4773-4778. DOI:10.1080/00207543.2021.1956675

Rakholia, R., Suárez-Cetrulo, A.L., Singh, M., & Carbajo, R.S. (2024). Advancing Manufacturing Through Artificial Intelligence: Current Landscape, Perspectives, Best Practices, Challenges, and Future Direction. *IEEE Access*, 12, 131621-131637. DOI: 10.1109/ACCESS.2024.3458830

Ramentol, E., Olsson, T., & Barua, S. (2021). Machine learning models for industrial applications. In K. G. Kyprianidis, K., & E. Dahlquist (Eds.), *AI and learning systems – Industrial applications and future directions* (Chap. 7). IntechOpen. DOI:10.5772/INTECHOPEN.93043

Salem, A.M., Yakoot, M.S., & Mahmoud, O. (2022). Addressing Diverse Petroleum Industry Problems Using Machine Learning Techniques: Literary Methodology—Spotlight on Predicting Well Integrity Failures. *ACS Omega*, 7(3), 2504-2519. DOI:10.1021/acsomega.1c05658

Srivastava, S., Pandey, V.K., Singh, K., Dar, A.H., Dash, K.K., Shams, R., Shaikh, A.M, & Kovács, B. (2024). Advances in Detection Technology for Authentication of Vegetable Oils: A Comprehensive Review. *Heliyon*, 10(15), e34759. DOI:10.1016/j.heliyon.2024.e34759

Verma, D. (2018). Analysis of Smart Manufacturing Technologies for Industry Using AI Methods. *Turkish Journal of Computer and Mathematics Education (TURCOMAT)*, *9*(2), 529-540. https://doi.org/10.17762/turcomat.v9i2.13857

Wang, H., Chen, Q., Zhao, J., Xu, L., Li, M., Zhao, Y., Zhao, Q., & Lu, Q. (2024). Dual-Modal Edible Oil Impurity Dataset for Weak Feature Detection. *Scientific Data*, 11(1), 1426. DOI:10.1038/s41597-024-04305-w