Exploration of Foods and Foodomics

Open Exploration

Open Access Original Article

Application of hyperspectral imaging techniques for sorting coffee beans

Antonio Berardi[1*], Karine Sophie Leheche Ouette[2], Alessandro Leone[1], Leonardo Feola[1], Cosimo Damiano Dellisanti[1], Domenico Tarantino[1], Antonia Tamborrino[1]

[1]Department of Agricultural and Environmental Science, University of Bari Aldo Moro, 70126 Bari, Italy

[2]Department of the Science of Agriculture, Food and Environment, University of Foggia, 71100 Foggia, Italy

***Correspondence:** Antonio Berardi, Department of Agricultural and Environmental Science, University of Bari Aldo Moro, Via Amendola 165/A, 70126 Bari, Italy. antonio.berardi@uniba.it

Academic Editor: Zhaowei Zhang, Oil Crops Research Institute of the Chinese Academy of Agricultural Sciences, China

Received: September 2, 2025 **Accepted:** April 22, 2026 **Published:** May 18, 2026

Cite this article: Berardi A, Ouette KSL, Leone A, Feola L, Dellisanti CD, Tarantino D, et al. Application of hyperspectral imaging techniques for sorting coffee beans. Explor Foods Foodomics. 2026;4:1010148. https://doi.org/10.37349/eff.2026.1010148

Abstract

Aim: Green coffee processing, before the roasting phase, requires effective removal of foreign materials and defective kernels to ensure product quality, process safety, and compliance with industrial requirements. The aim of this research is to use conventional RGB-based optical sorters for product sorting. These rely primarily on surface colour characteristics and can be limited when contaminants display visual similarities to healthy beans.

Methods: Hyperspectral imaging (HSI) provides a non-destructive alternative by integrating spatial and spectral information in the visible and near-infrared (VIS/NIR) range. In this study, a VIS/NIR HSI system was integrated into a commercial industrial optical sorter and validated under real operating conditions. Contaminated green coffee batches (10 kg) containing known amounts of organic and inorganic contaminants were processed through multiple sorting passes using a statistical classification logic embedded into the sorter programmable logic controller (PLC) for real-time decision making.

Results: The system achieved complete removal of stone contaminants after a single pass, while organic contaminants (peel and defective beans) were substantially reduced across successive cycles. After two sorting passes, the cumulative yield of compliant coffee beans was approximately 84%, representing an acceptable trade-off between contaminant removal efficiency and product loss in an industrial context.

Conclusions: Overall, the results support the feasibility of deploying VIS/NIR hyperspectral sensing for high-throughput industrial coffee sorting, with potential advantages in discrimination capability compared with conventional colour-based systems.

Keywords

hyperspectral imaging, VIS/NIR, industrial optical sorter, coffee, real-time classification

Introduction

Coffee is a globally important product mainly cultivated in regions such as South America, Africa, and Southeast Asia. The total production of coffee exceeds 7 million tons annually [1]. Coffee is a globally traded commodity with a complex supply chain that includes harvesting, post-harvest processing, storage, transport, roasting, and grinding. Prior to roasting, green coffee beans typically undergo cleaning and sorting steps to remove foreign materials and non-compliant fractions that can compromise downstream processing and final product quality. Defects and contaminants such as stones, husks, broken beans, and visually defective kernels may negatively affect roasting performance, cause mechanical damage to equipment, and deteriorate sensory characteristics of the beverage [2].

Before roasting, the selection of coffee beans is a fundamental step to ensure the final product is of high quality. This selection process involves both mechanical and manual methods.

From a sustainability perspective, optimizing the sorting process helps reduce food waste and resource loss in coffee value chains. Improved sorting also minimizes the need for redundant processing stages and contributes to a more efficient, low-impact production system that aligns with global goals for sustainable agriculture and industrial processes.

Mechanically, systems are used to classify beans based on size, shape, and to remove foreign materials like stones, metal fragments, and other contaminants [3]. These systems help streamline the process and improve consistency. Manual selection, on the other hand, involves human operators inspecting the beans to eliminate defective or damaged ones, ensuring only the best beans proceed to roasting.

Industrial sorting is traditionally performed using mechanical separation coupled with optical systems. For example, artificial vision systems analyze surface features like colour, shape, size, and visible defects to identify and remove defective beans. These systems are quite effective in external defect detection, but have limitations; they cannot detect internal defects or contaminants hidden inside the beans [4]. Their effectiveness improves when the final product is more homogeneous, but they still serve as partial solutions. RGB-based optical sorters classify products mainly by surface colour, size, and shape. While effective for many visible defects, RGB approaches may be less reliable when contaminants share a similar visual appearance with sound beans or when discrimination would benefit from information linked to chemical composition rather than colour alone [5, 6].

In recent years, the food industry has been increasingly adopting automated process control systems to ensure consistent quality and reduce processing times [7–10].

Hyperspectral imaging (HSI) has emerged as a powerful non-destructive technique for food quality and safety assessment because it combines imaging and spectroscopy to acquire spatially resolved spectra over multiple wavelengths, often in the visible and near-infrared (VIS/NIR) region [11–13]. HSI has been extensively investigated for classification, grading, and defect detection across many agri-food products, demonstrating that spectral signatures can enhance discrimination beyond conventional imaging [11–13]. In the coffee sector, spectroscopic approaches (including NIR) have been used for quality evaluation and defect-related analysis, largely under laboratory or offline conditions [14, 15].

Despite this progress, transferring hyperspectral methods from laboratory studies to industrial environments [16] remains challenging due to constraints on acquisition speed, data processing, real-time classification, and integration within existing process-control architectures. Industrial deployment requires robust and computationally efficient models that can operate at high throughput and interface with programmable logic controller (PLC)-controlled rejection systems. Many studies have shown how useful it is for different products like vegetables [17], oil extraction [18], cereal processing [19, 20], and legumes [18]. Plus, it's quite cost-effective to manage [21–26].

The novelty of the present study lies in the industrial-scale validation of a VIS/NIR HSI system integrated into a commercial optical sorter for green coffee processing. Rather than introducing a new sensor design, the manuscript focuses on feasibility and performance under real operating conditions, quantifying contaminant removal and yield across sorting passes. The specific objectives are to: (i) evaluate discrimination of green coffee beans from typical organic and inorganic contaminants using VIS/NIR hyperspectral sensing; (ii) quantify sorting performance (contaminant removal, product loss, yield) under industrial conditions; and (iii) discuss industrial implications and positioning relative to conventional colour sorting.

Materials and methods

The classification models were implemented on an industrial optical calibration machine equipped with VIS/NIR sensors in the wavelength range 400–1,700 nm in order to discriminate the contaminants present in the coffee samples, according to the experimental plan prepared.

Spectral data processing and model development were performed offline in MATLAB® (MathWorks, Natick, MA, USA). Standard hyperspectral preprocessing procedures (commonly used in VIS/NIR food analysis) were applied to reduce acquisition variability and improve robustness, including dark/white reference correction and spectral normalization [14, 15].

The classification strategy adopted in this work is a statistical/chemometric approach, selected for computational efficiency, interpretability, and compatibility with real-time industrial constraints. Statistical classifiers remain widely used for hyperspectral food applications and for industrial implementations where robustness and fast inference are required [14, 15]. Although deep learning methods have become increasingly prominent in hyperspectral analysis, their deployment in real-time industrial sorting can be limited by computational load and integration constraints, particularly when strict timing and PLC interfacing are required [27, 28]. For these reasons, a statistical classifier was preferred for the present industrial validation.

The classifier was trained using representative samples of sound coffee beans and each contaminant class (peel, defective beans, stones). Following offline verification, the resulting decision logic was translated into a PLC-compatible set of classification rules and embedded into the sorter control architecture for real-time rejection.

Raw material

The raw material consisted of green coffee beans (*Coffea arabica*), Figure 1, imported from South America. Prior to the trials, homogeneous batches were prepared to ensure repeatability. The incoming material was manually inspected to identify and separate the main categories relevant to industrial quality control (sound beans, defective beans, and foreign materials) [2].

The contaminants considered in this study were: (i) coffee husks/peel; (ii) visually defective beans (hereafter "coffee dotted" as in the original dataset); and (iii) stones. These categories reflect common non-compliant fractions encountered in green coffee processing and cleaning stages [2, 14].

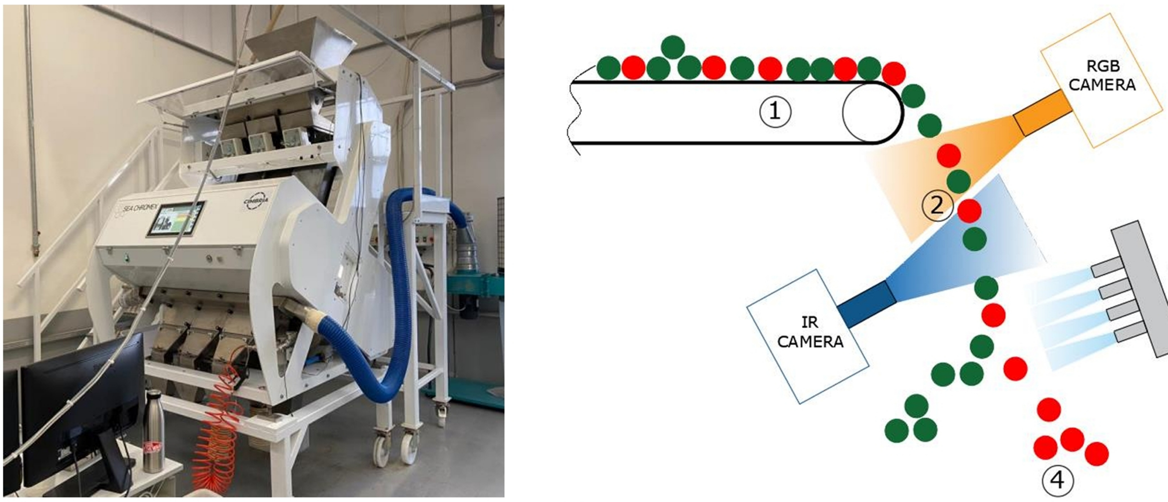
Each category was counted (when applicable) and weighed. Known quantities of contaminants were then manually separated and reintroduced into coffee batches to obtain controlled contamination levels suitable for quantitative evaluation of sorting performance, consistent with validation approaches for optical inspection systems.

Industrial hyperspectral sorter

Trials were carried out using a commercial vertical optical sorter (Figure 2a) configured with multi-sensor inspection. The machine is composed of independent gravity-fed channels. Each channel includes two NIR line-scan cameras positioned on opposite sides of the product flow and two RGB cameras (front and rear) operating in the visible spectrum (Figure 2b). Hyperspectral/Line-scan configurations are widely adopted for high-throughput food inspection due to their suitability for conveyor or free-fall-based operations.

The hyperspectral inspection covered the VIS/NIR region (nominal range 400–1,700 nm as specified by the instrument configuration). The system operated with scan rates up to 15,000 Hz and optical spatial resolution of approximately 0.06 mm (60 μm), enabling industrial throughput. In-line hyperspectral applications require careful balancing of acquisition speed and classification complexity to ensure real-time feasibility.

Figure 1. Coffee beans for test.



(a) (b)

Figure 2. Automatic industrial sorter: (a) photograph of the equipment; (b) schematic diagram of the system. ① vibrating plate, ② inclined slide, ③ compressed air jets, ④ product unloading.

Classification output triggered rejection by compressed air jets controlled via solenoid valves. The sorter was managed by a PLC; the classification logic (derived offline) was implemented in the PLC-compatible environment for real-time operation.

Experimental procedure

The tests, in accordance with the experimental plan, were performed with homogeneous batches of sound coffee beans.

Each experimental batch had a nominal mass of 10 kg to which contaminants were added as reported in Table 1. Each test was repeated 5 times and 3 consecutive passages of the product inside the sorter to evaluate performance as a function of sorting cycles (multi-pass strategies are commonly used to refine optical separation when contaminants have partial overlap with compliant product features).

Table 1. Sample composition of coffee contaminated.

Description	Quantity (g)	Percentage (%)	Number of units
Coffee no defects	10,000.00	96.90	-
Peel	70.00	0.68	763
Coffee dotted	215.00	2.08	1,523
Stone	35.12	0.34	33
Total	**10,320.12**	**100.00**	-

After each pass, accepted and rejected fractions were collected separately. Each fraction was weighed and manually inspected to determine the mass (and units, when relevant) of each category.

Results

First sorting pass

The composition of the rejected fraction after the first pass is shown in Table 2. Stones were completely removed during the first selection cycle (100% removal, 33/33 units), demonstrating the effectiveness of the system in eliminating dense foreign materials that can damage downstream equipment.

Table 2. Sample composition of coffee discarded—first selection.

Description	Quantity (g)	Percentage of sample (%)	Number of units	Percentage of contaminant (%)
Coffee no defects	988.38	9.88	-	-
Peel	59.81	0.60	652	85.45
Coffee dotted	146.15	1.46	1,036	68.02
Stone	35.12	0.35	33	100.00
Total	**1,229.46**	**12.29**	-	-

A substantial fraction of organic contaminants was also removed after the first pass: peel removal reached 85.45% (652/763 units), and defective beans ("coffee dotted") removal reached 68.02% (1036/1523 units). The first pass also rejected a fraction of sound coffee beans (988.38 g), corresponding to 9.88% of the initial 10 kg baseline.

Second sorting pass

The rejected fraction after the second pass is reported in Table 3. As expected, no stones were detected in the second-pass rejects, confirming complete elimination in the first pass. Residual organic contaminants were further reduced, with additional removal of peel and defective beans.

Table 3. Sample composition of coffee discarded—second selection.

Description	Quantity (g)	Percentage of sample (%)	Number of units	Percentage of contaminant (%)
Coffee no defects	562.31	5.62	-	-
Peel	8.34	0.08	91	11.93
Coffee dotted	47.41	0.47	336	22.06
Stone	0.00	0.00	0	0.00
Total	**618.06**	**6.17**	-	-

In the second pass, the mass of sound coffee rejected (562.31 g) was lower than in the first pass, while contaminant-related rejects were reduced to small residual quantities.

Overall yield

The cumulative product loss in terms of sound coffee rejected was 988.38 g after the first pass and 562.31 g after the second pass. Therefore, the overall yield of compliant coffee after two sorting passes was approximately 84% of the initial 10 kg sound coffee baseline, consistent with the yield trend reported in the original manuscript.

Discussion

This study provides an industrial-scale validation of VIS/NIR hyperspectral sorting applied to green coffee, addressing a key gap between laboratory demonstrations and real-time deployment. The complete removal of stones after the first pass is particularly relevant for industrial processing, as foreign inorganic materials can cause mechanical damage and safety issues. Compared with conventional RGB-based sorting, which relies mainly on surface colour information, hyperspectral sensing can exploit spectral features related to material composition, potentially improving discrimination when visual appearance overlaps.

The progressive reduction of organic contaminants (peel and defective beans) across sorting passes indicates that the chosen multi-pass strategy is effective for refining product quality (Figure 3). Figure 4a and 4b show the results of the selection, respectively, the selected coffee product and the rejected product with the contaminants. Multi-pass operation is consistent with industrial practice for optical sorting systems when the objective is to maximize contaminant removal while balancing product retention. In our case, most contaminants were removed in the first pass, while the second pass primarily acted as a "polishing" step, further reducing residual fractions.

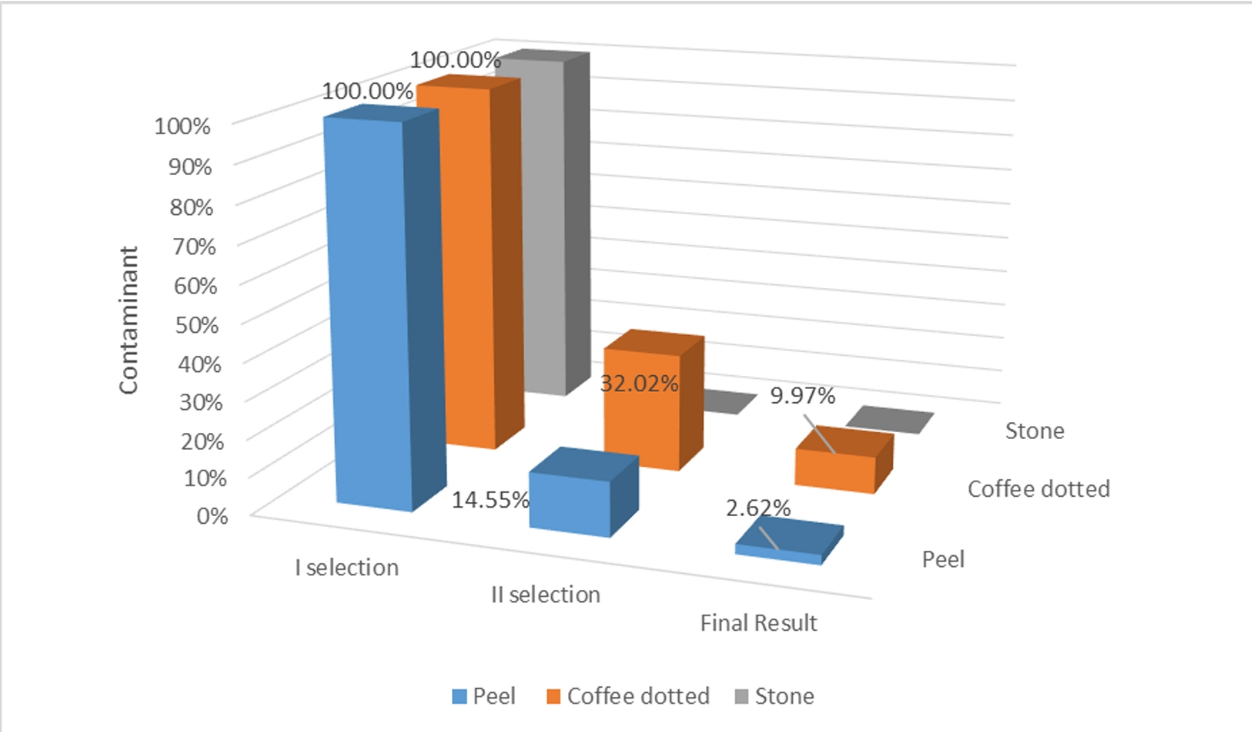


Figure 3. Residual contaminants for coffee discrimination.

A key industrial trade-off is the loss of a compliant product. The cumulative yield of approximately 84% reflects the balance between strict rejection criteria and product retention (Figure 5). In high-value processing lines, moderate product losses can be acceptable when removal of harmful contaminants and improvement of batch uniformity reduce downstream risks and enhance final quality. Importantly, the primary contribution of this work is not proposing a new algorithm but demonstrating that a computationally efficient statistical classifier can be embedded into a PLC-based sorter architecture and operate in real time under industrial constraints. This positioning is aligned with the broader literature, highlighting that industrial HSI adoption depends strongly on robust preprocessing, efficient models, and system integration rather than accuracy alone.

While deep learning approaches are increasingly reported for hyperspectral analysis and have shown strong performance in several food applications, their use in industrial sorting remains constrained by computational requirements, model deployment complexity, and the need for deterministic real-time

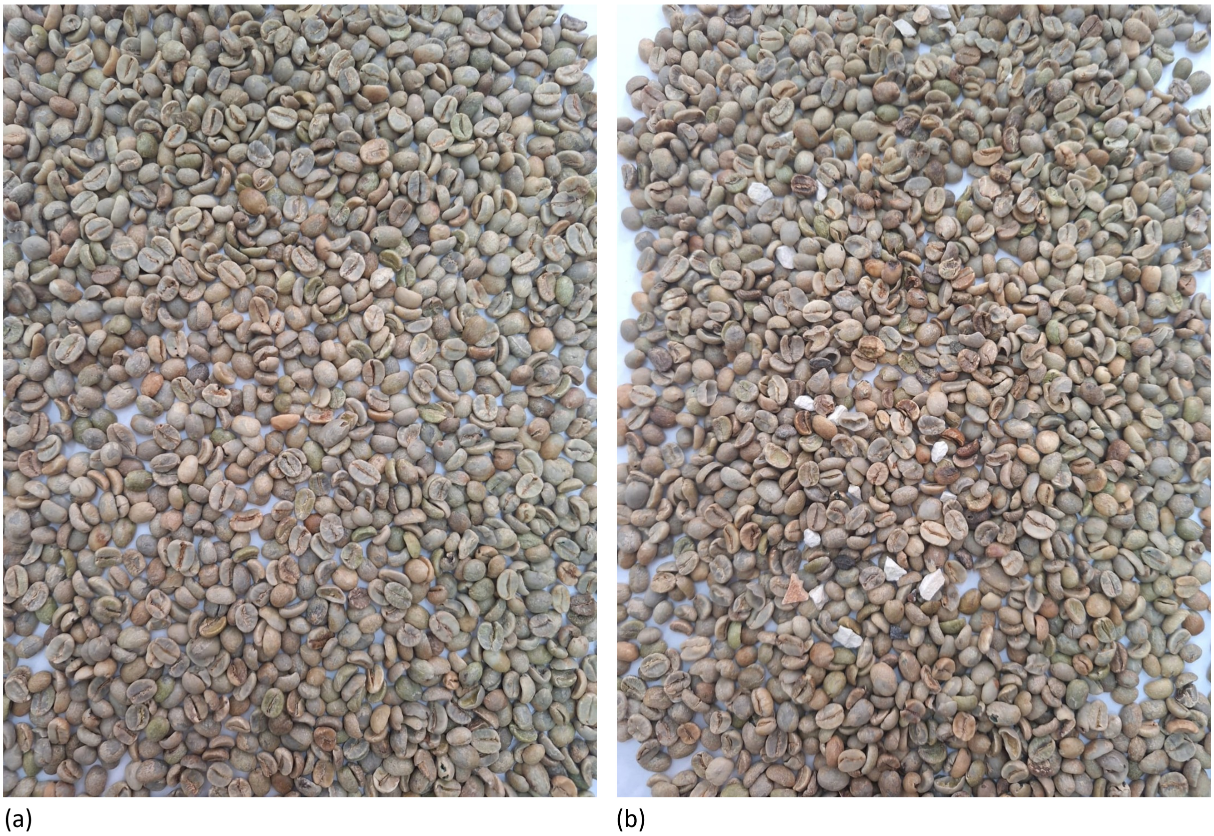


Figure 4. Representative images of coffee beans after sorting, showing selected (a) and discarded (b).

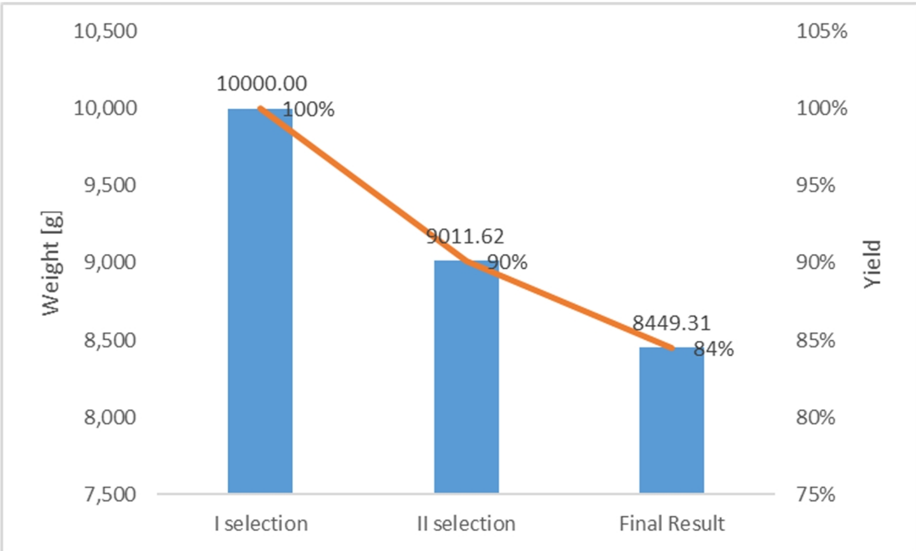


Figure 5. Optical sorting performance on the selected product.

execution. In this context, the statistical approach adopted here represents a pragmatic solution for industrial deployment, while deep learning remains a promising direction for future upgrades when hardware/software constraints permit.

From a sustainability perspective, improved sorting accuracy can reduce waste and enhance resource utilization by limiting reprocessing and preventing the progression of contaminated material into downstream steps. More broadly, advanced physical technologies and automation are increasingly recognized as key enablers of sustainable food processing chains.

Conclusions

This work validated the industrial deployment of a VIS/NIR HSI system integrated into a commercial optical sorter for green coffee processing. The system achieved complete removal of stone contaminants after a single pass and substantial reduction of organic contaminants across successive sorting cycles. After two sorting passes, the cumulative yield of compliant coffee beans was approximately 84%, representing an acceptable industrial trade-off between product retention and contaminant removal.

The results support the feasibility of real-time hyperspectral sorting in high-throughput environments and highlight the added value of spectral information compared with conventional colour-based systems. Overall, HSI represents a robust and non-destructive tool for advanced industrial quality control in coffee processing, with potential extension to other agri-food sectors. Future work should focus on optimizing classification rules to improve yield, expanding defect categories, and evaluating advanced machine learning approaches under industrial deployment constraints.

Abbreviations

HSI: hyperspectral imaging

PLC: programmable logic controller

VIS/NIR: visible and near-infrared

Declarations

Author contributions

AB: Conceptualization, Writing—original draft. KSLO: Investigation. AL: Supervision. LF: Investigation. CDD: Investigation. DT: Conceptualization. AT: Writing—review & editing. All authors read and approved the submitted version.

Conflicts of interest

The authors declare that they have no conflicts of interest.

Ethical approval

Not applicable.

Consent to participate

Not applicable.

Consent to publication

Not applicable.

Availability of data and materials

The raw data supporting the conclusions of this manuscript will be made available by the authors, without undue reservation, to any qualified researcher.

Funding

Not applicable.

Publisher's note

Open Exploration maintains a neutral stance on jurisdictional claims in published institutional affiliations and maps. All opinions expressed in this article are the personal views of the author(s) and do not represent the stance of the editorial team or the publisher.

References

1. European Coffee Report 2023-2024 [Internet]. ECF; c2026 [cited 2025 Sep 2]. Available from: https://www.ecf-coffee.org/european-coffee-report-2023-2024/
2. Farah A. Coffee Constituents. In: Chu YF, editor. Coffee: Emerging Health Effects and Disease Prevention. Hoboken: Wiley-Blackwell; 2012.
3. Cheremisinoff NP. Chapter 6 Mechanical Separation Equipment. In: Handbook of Chemical Processing Equipment. Amsterdam: Elsevier; 2000. pp. 334–434.
4. Ariana DP, Lu R. Quality evaluation of pickling cucumbers using hyperspectral reflectance and transmittance imaging: Part I. Development of a prototype. Sens Instrumen Food Qual. 2008;2: 144–51. [DOI]
5. Tadhg B, Da-Wen S. Improving quality inspection of food products by computer vision––a review. J Food Eng. 2004;61:3–16. [DOI]
6. Blasco J, Aleixos N, Moltó E. Machine Vision System for Automatic Quality Grading of Fruit. Biosyst Eng. 2003;85:415–23. [DOI]
7. Juliano P, Gaber MAFM, Romaniello R, Tamborrino A, Berardi A, Leone A. Advances in Physical Technologies to Improve Virgin Olive Oil Extraction Efficiency in High-Throughput Production Plants. Food Eng Rev. 2023;15:625–42. [DOI]
8. Perone C, Romaniello R, Leone A, Catalano P, Tamborrino A. CFD Analysis of a Tubular Heat Exchanger for the Conditioning of Olive Paste. Appl Sci. 2021;11:1858. [DOI]
9. Romaniello R, Perone C, Tamborrino A, Berardi A, Leone A, Di Taranto A, et al. Additives Individuation in Raw Ham Using Image Analysis. Chem Eng Trans. 2021;87:217–22. [DOI]
10. Tamborrino A, Catalano F, Berardi A, Bianchi B. New modelling approach for the energy and steam consumption evaluation in a fresh pasta industry. Chem Eng Trans. 2021;87:409–14. [DOI]
11. Huang H, Liu L, Ngadi MO. Recent developments in hyperspectral imaging for assessment of food quality and safety. Sens (Basel). 2014;14:7248–76. [DOI] [PubMed] [PMC]
12. Qin J, Kim M, Chao K, Chan D, Delwiche S, Cho B. Line-Scan Hyperspectral Imaging Techniques for Food Safety and Quality Applications. Appl Sci. 2017;7:125. [DOI]
13. Wu D, Sun D. Advanced applications of hyperspectral imaging technology for food quality and safety analysis and assessment: A review — Part II: Applications. Innov Food Sci Emerg Technol. 2013;19: 15–28. [DOI: 10.1016/j.ifset.2013.04.016].
14. Santos JR, Sarraguça MC, Rangel AOSS, Lopes JA. Evaluation of green coffee beans quality using near infrared spectroscopy: a quantitative approach. Food Chem. 2012;135:1828–35. [DOI] [PubMed]
15. Dias RCE, Valderrama P, Março PH, Dos Santos Scholz MB, Edelmann M, Yeretzian C. Data on roasted coffee with specific defects analyzed by infrared-photoacoustic spectroscopy and chemometrics. Data Brief. 2018;20:242–9. [DOI] [PubMed] [PMC]
16. Jiang Z, Lv A, Zhong L, Yang J, Xu X, Li Y, et al. Rapid Prediction of Adulteration Content in *Atractylodis rhizoma* Based on Data and Image Features Fusions from Near-Infrared Spectroscopy and Hyperspectral Imaging Techniques. Foods. 2023;12:2904. [DOI] [PubMed] [PMC]
17. Amodio M, Berardi A, Ricci I, Babellahi F, Colelli G. Use of hyperspectral imaging for the discrimination of artichoke by cultivar and harvest time. Acta Hortic. 2020;1284:165–72. [DOI]
18. Leone A, Berardi A, Antonelli G, Dellisanti CD, Tamborrino A. NIR Spectroscopy for the Online Monitoring of Water and Olive Oil Content in Pomace during the Extraction Process. ASI. 2024;7:96. [DOI]
19. Romaniello R, Barrasso AE, Berardi A, Perone C, Tamborrino A, Catalano F, et al. Hyperspectral imaging system to on-line monitoring the soy flour content in a functional pasta. J Agric Eng. 2023;54: 1535. [DOI]

20. Romaniello R, Barrasso AE, Perone C, Tamborrino A, Berardi A, Leone A. Optimisation of an Industrial Optical Sorter of Legumes for Gluten-Free Production Using Hyperspectral Imaging Techniques. Foods. 2024;13:404. [DOI] [PubMed] [PMC]
21. Kılıç K, Boyacı İH, Köksel H, Küsmenoğlu İ. A classification system for beans using computer vision system and artificial neural networks. J Food Eng. 2007;78:897–904. [DOI]
22. Beghi R, Giovenzana V, Civelli R, Cini E, Guidetti R. Characterisation of olive fruit for the milling process by using visible/near infrared spectroscopy. J Agricult Engineer. 2013;44:8. [DOI]
23. Lee H, Cho B, Kim MS, Lee W, Tewari J, Bae H, et al. Prediction of crude protein and oil content of soybeans using Raman spectroscopy. Sens Actuators B: Chem. 2013;185:694–700. [DOI]
24. Mendoza FA, Cichy KA, Sprague C, Goffnett A, Lu R, Kelly JD. Prediction of canned black bean texture (Phaseolus vulgaris L.) from intact dry seeds using visible/near infrared spectroscopy and hyperspectral imaging data. J Sci Food Agric. 2018;98:283–90. [DOI] [PubMed]
25. Wei X, Zheng W, Zhu S, Zhou S, Wu W, Xie Z. Application of terahertz spectrum and interval partial least squares method in the identification of genetically modified soybeans. Spectrochim Acta A Mol Biomol Spectrosc. 2020;238:118453. [DOI] [PubMed]
26. Perone C, Romaniello R, Leone A, Berardi A, Catalano P, Tamborrino A. CFD Analysis of a Tube-in-tube Heat Exchanger to Pre-heat Olive Pastes. Chem Eng Trans. 2021;87:253–8. [DOI]
27. Das M, Wan SY, Agus S. A review of machine learning in hyperspectral imaging for food safety. Vib Spectrosc. 2025;139:103828. [DOI]
28. Yang C, Guo Z, Barbin DF, Dai Z, Watson N, Povey M, et al. Hyperspectral Imaging and Deep Learning for Quality and Safety Inspection of Fruits and Vegetables: A Review. J Agric Food Chem. 2025;73: 10019–35. [DOI] [PubMed]