

Systematic Review

Computer Vision and XRF-IoT Sensor Systems for Detecting Heavy Metals in Export Crops: A Comprehensive Systematic Review

Kevin Tupac-Agüero ¹, Kenneth Ortega-Moran ¹, Javier Gamboa-Cruzado ^{1,*} , Rosa Menéndez Mueras ¹, Carlos Del-Valle-Jurado ², Alex Salazar-Marzal ³ and Angel Nuñez Meza ⁴

¹ Facultad de Ingeniería de Sistemas e Informática, Universidad Nacional Mayor de San Marcos, Lima 15081, Peru; kevin.tupac@unmsm.edu.pe (K.T.-A.); ekenneth.ortega@unmsm.edu.pe (K.O.-M.); rmenendezm@unmsm.edu.pe (R.M.M.)

² Facultad de Ingeniería Geológica, Minera, Metalúrgica y Geográfica, Universidad Nacional Mayor de San Marcos, Lima 15081, Peru; cdelvallej@unmsm.edu.pe

³ Facultad de Ciencias Administrativas, Universidad Nacional Mayor de San Marcos, Lima 15081, Peru; alexmelecio.salazar@unmsm.edu.pe

⁴ Facultad de Ingeniería de Sistemas, Universidad Nacional Daniel Alcides Carrión, Pasco 19001, Peru; anunezm@undac.edu.pe

* Correspondence: jgamboac@unmsm.edu.pe; Tel.: +51-999117404

Abstract

The increasing concern over heavy metal contamination in export crops has intensified research on the application of computer vision systems (CVS) and advanced sensing technologies within multi-level agricultural monitoring frameworks spanning soil contamination assessment, crop spectral diagnostics, and in situ elemental sensing. This study conducts a systematic literature review following Kitchenham's methodology, from which 68 studies were finally included after screening and eligibility assessment. The review focuses on the use of hyperspectral imaging (HSI) and XRF-IoT sensors (X-ray fluorescence units enhanced with IoT connectivity) for detecting heavy metals in export crops, considering publications from the last seven years indexed in Web of Science Core Collection, Scopus, IEEE Xplore, EBSCOhost, and Springer Nature Link. The findings indicate that research is concentrated in highly digitalized countries, which limits its global applicability; moreover, a substantial proportion of studies is published in Q1 journals, although the methodologies are not always fully objective. Likewise, the most developed research lines are oriented toward image-based diagnostics and crop analysis. These results reveal a gap between technological advances in computer vision and their integration into agricultural decision-making aimed at improving the quality of export crops. It is recommended to foster research with greater geographical diversity, grounded in solid theoretical frameworks and an ethical perspective.



Academic Editors: Panayiotis Bozanis and Leonidas Akritidis

Received: 24 January 2026

Revised: 22 February 2026

Accepted: 23 February 2026

Published: 26 February 2026

Copyright: © 2026 by the authors.

Licensee MDPI, Basel, Switzerland.

This article is an open access article distributed under the terms and conditions of the [Creative Commons Attribution \(CC BY\)](https://creativecommons.org/licenses/by/4.0/) license.

Keywords: computer vision; hyperspectral imaging; IoT sensor; X-ray fluorescence (XRF); heavy metals; agro-exportation; export crops; systematic literature review

1. Introduction

The safety and quality of export crops constitute fundamental pillars for international trade and public health protection. In this context, emerging technologies such as hyperspectral imaging (HSI), the Internet of Things (IoT), and Artificial Intelligence (AI) are transforming precision agriculture by enabling advanced monitoring and the early

detection of contaminants and phytosanitary risks. These innovations support detailed, real-time assessment of soil and crop conditions, thereby facilitating more efficient resource management and compliance with international food safety regulations.

In recent years, the detection of contaminants in agricultural systems has undergone a radical transformation due to the convergence of computer vision, deep learning, and hyperspectral image processing. Studies such as [1–3] demonstrate HSI's efficacy in mineral classification through CNN–LSTM hybrid models, while works such as [4–6] validate portable sensing technologies for field deployment. Likewise, studies such as [7–9] show that advanced architectures—including DODTL-CTDC and 3-D UNet models—enable highly reliable agricultural mapping by optimizing hyperparameters and exploiting the spectral richness of hyperspectral data cubes.

Parallel developments in hyperspectral and remote-sensing technologies have expanded toward broader environmental monitoring challenges in agriculture. The authors [10–12] demonstrate that spectral reflectance, radar imagery, and machine-learning models such as BPANN and Random Forest can accurately estimate soil and water-related variables. Although these studies focus on aspects such as water stress, salinity, or hydraulic conductivity, their methodological relevance lies in demonstrating how imaging-based techniques can detect and quantify contaminants in agricultural systems. Similarly, the studies [13–15] highlight how predictive tools—including pesticide-risk models and hybrid environmental frameworks—identify contamination hotspots affecting millions of people, reinforcing the potential of integrated sensor platforms to detect pollutants such as heavy metals in soil and water.

Moreover, advances in multimodal data fusion have strengthened early detection capabilities in agricultural systems. The authors [16–18] show that architectures such as DeepLabV3+ and LIDAR–electro-optical systems achieve semantic segmentation accuracies above 84% in vineyards and large-scale crops. Complementarily, the works [19–21] illustrate how LSTM models and long-term time-series analyses help characterize pollutant-related dynamics—such as N₂O emissions or copper accumulation in soils—by identifying key environmental predictors including precipitation, organic carbon, and nitrates. More systemic convergence frameworks integrating IoT, remote sensing, Agriculture 4.0, generative AI, and edge computing have also been proposed to enable real-time monitoring of high-value crops [22–24].

Finally, studies such as [25–27] emphasize that beyond technical accuracy, the real impact of these technologies depends on systemic sustainability, including fertilizer management, declining selenium deposition, and the need for policies regulating metal contamination in agricultural soils. Thus, the current literature advances not only in predictive modeling but also in holistic frameworks linking technical precision with environmental sustainability and global equity.

1.1. Research Gap

Despite advances in hyperspectral imaging, deep learning, IoT integration, and multimodal sensing frameworks, significant limitations persist. These include the limited availability of public datasets for heavy-metal detection, high implementation costs, insufficient farmer training, the scarcity of economic evaluations of hybrid HSI + XRF-IoT platforms, model generalization challenges, and the absence of standardized replicability frameworks. While individual studies have validated imaging systems, portable sensors, or AI-based models independently, their systematic integration for heavy-metal detection in export crops remains insufficiently synthesized and critically evaluated within a unified monitoring framework in the literature.

1.2. Motivation

Given the strategic importance of export crops for global trade, food security, and environmental sustainability, a structured synthesis is required to consolidate advances in computer vision, hyperspectral imaging, and XRF-IoT sensing within a unified monitoring architecture. Such synthesis should evaluate not only technical performance but also methodological robustness, economic feasibility, scalability, and replicability to support sustainable and transferable solutions aligned with international regulatory standards.

Accordingly, this study conducts a comprehensive systematic literature review to critically synthesize advances, limitations, and research opportunities in Computer Vision Systems for detecting heavy metals in export crops using hyperspectral imaging and XRF-IoT sensors. The review aims to provide an analytical framework capable of supporting future research directions and strengthening technology-driven agricultural monitoring strategies.

The remainder of this paper is organized as follows: Section 2 presents the theoretical framework underlying the study; Section 3 details the methodology applied in the review process; Section 4 presents the main results along with their discussion; and Section 5 provides the conclusions, highlighting the implications of the study and formulating recommendations for future research directions.

2. Theoretical Background

Due to recent advances in the application of computer vision systems for detecting heavy metals in export crops through hyperspectral imaging and XRF-IoT sensors, it is essential to establish a rigorous and coherent conceptual foundation that allows for a clear understanding of the theoretical, methodological, and technological principles before analyzing in depth the trends that currently shape and guide this field of research.

2.1. Computer Vision Systems

Computer vision systems represent a key technology in precision agriculture, enabling machines to interpret and analyze real-world imagery through advanced deep-learning architectures that emulate human visual perception [28,29]. Advances in image processing and machine learning have allowed computer vision systems to become increasingly accurate and effective in the classification of agricultural images [30]. Integration with the Internet of Things (IoT) has facilitated the real-time collection of critical variables, such as soil moisture and temperature, thereby optimizing the use of resources such as water and fertilizers [31]. These systems are capable of analyzing large volumes of visual data without compromising crop quality or development, making them ideal non-destructive tools for agriculture [32]. Moreover, the incorporation of deep learning and convolutional neural networks (CNN) has enhanced the ability to identify complex patterns and make more precise predictions regarding crop status [33]. By merging data obtained through computer vision with complementary sensors such as moisture or temperature devices, farmers can access more detailed and reliable information on crop conditions [34]. Nevertheless, despite the progress of these techniques, challenges persist, such as the need for adequate model training and the availability of high-quality data, factors that limit widespread adoption [35]. Artificial Intelligence (AI) also constitutes an essential tool for optimizing and automating mechanical tasks that require the prior analysis of large datasets [36]. Finally, AI aims to emulate human cognitive abilities, enabling machines to perform tasks such as learning, reasoning, and self-correction [37].

In this review, hyperspectral imaging (HSI) is conceptualized as an advanced spectral modality within computer vision systems, since it relies on image acquisition, spectral-spatial feature extraction, and AI-based modeling pipelines analogous to those used in

conventional RGB and multispectral computer vision. Therefore, HSI-based soil contamination mapping and crop-level diagnostics are treated as integral components of computer-vision-driven agricultural monitoring frameworks rather than as separate thematic domains.

Within the reviewed literature, Unmanned Aerial Vehicles (UAVs) equipped with RGB cameras offer faster and more cost-effective deployment than depth cameras or laser scanners, although they face constraints such as limited battery life and flight regulations [38]. In this review, UAVs are treated as modular carriers capable of integrating RGB, hyperspectral, or multimodal sensors according to the monitoring objective.

In parallel, XRF-IoT sensors are critical for direct, non-destructive, and element-specific quantification of heavy metals (e.g., Pb, Cd, As), and when integrated with IoT communication modules, they enable real-time monitoring and wireless reporting of contamination levels in agricultural environments. Together, these technologies form a complementary diagnostic framework in which HSI provides spectral indicators of contamination while XRF-IoT offers precise elemental measurements, making them fundamental for reliable heavy-metal detection in export crops.

2.2. Detection of Heavy Metals in Export Crops Using Hyperspectral Imaging and XRF-IoT Sensors

The detection of heavy metals in export crops is essential to ensure food safety and comply with international regulatory standards [39]. Contaminants such as lead, cadmium, and arsenic can accumulate in agricultural soils, alter plant physiology, and compromise the quality of export products [40]. Hyperspectral imaging (HSI) has emerged as a powerful tool for identifying these contaminants because it captures hundreds of contiguous spectral bands, enabling the detection of subtle biochemical variations associated with heavy-metal stress and spectral anomalies not visible in RGB or multispectral imagery [41]. In parallel, XRF-IoT sensors—based on X-ray fluorescence spectroscopy integrated with IoT communication modules—provide non-destructive, element-specific quantification of heavy metals in real time, without the need for physical sample destruction or laboratory processing [42]. The combination of HSI and XRF-IoT thus offers complementary diagnostic capabilities: while HSI detects indirect spectral signatures in plant tissues, XRF-IoT provides direct chemical confirmation of contaminant concentration levels, strengthening the reliability and interpretability of agricultural monitoring systems [43].

However, the practical implementation of these technologies presents several challenges. HSI measurements are sensitive to environmental variation, illumination changes, and sensor calibration issues that can introduce spectral noise and reduce model generalization [44]. Likewise, XRF-IoT systems can suffer from instrument drift, matrix effects, and varying surface conditions that affect measurement stability and accuracy. Both technologies also depend on controlled sample preparation protocols, which are not always feasible in field environments.

Despite these limitations, the integration of hyperspectral imaging and XRF-IoT sensors represents a substantial advancement in precision agriculture. Their combined use enables continuous, non-invasive monitoring that improves early contaminant detection, supports timely decision-making, and strengthens compliance with international food safety requirements [45]. Ongoing research increasingly emphasizes data fusion strategies that merge spectral information with IoT-based measurements to enhance classification accuracy and reduce false positives [46].

Finally, scaling these technologies requires efforts in low-cost hardware adaptation, open-access datasets for model training, and farmer-centered capacity-building initiatives that facilitate adoption in diverse agricultural contexts [47]. When deployed effectively, hyperspectral imaging and XRF-IoT sensor systems have the potential to transform heavy-

metal monitoring in export crops, contributing to safer, traceable, and higher-quality agricultural production [48].

3. Review Method

This research is structured as a Systematic Literature Review (SLR), a rigorous and reproducible approach designed to identify, evaluate, and synthesize the available evidence related to a specific research question. To ensure methodological robustness and transparency, the framework proposed by Kitchenham et al. (2009) [49], widely recognized in the academic community, was adopted and aligned with the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) 2020 guidelines [50]. The systematic review process was organized into three interconnected phases, each with defined objectives and activities contributing to the overall purpose of the study: (1) review planning, which establishes the foundations and protocol of the review; (2) review execution, which includes database searching, study selection, and data extraction; and (3) results presentation, in which the findings are synthesized and reported in a structured manner. This systematic procedure helps reduce researcher bias and enhances the reliability and reproducibility of the results, which are essential for scientific validity.

The applied methodology aims to provide a comprehensive and up-to-date overview of the state of knowledge in computer vision systems for detecting heavy metals in export crops using hyperspectral technologies and XRF-IoT sensors. This systematic review did not include a preregistered protocol, as protocol registration is not common practice in engineering-oriented SLRs. Nevertheless, all methodological decisions were defined a priori—including the search strategy, screening criteria, and quality-assessment procedures—to ensure transparency and minimize bias. Full compliance with PRISMA 2020 is documented in Appendix A (PRISMA 2020 Checklist).

3.1. Research Questions

The research questions (RQs) are formulated with the purpose of providing specific answers, which grants this systematic review paper greater relevance by selecting aspects that require further exploration. The questions are as follows:

- RQ1. What performance indicators are used to measure the effectiveness of Computer Vision Systems?
- RQ2. Which programming languages are most frequently used in the development of Computer Vision Systems?
- RQ3. What are the main conceptual trends in research on Computer Vision Systems and their impact on the Detection of Heavy Metals in Export Crops using Hyperspectral Imaging and XRF-IoT Sensors?
- RQ4. How are studies on the application of Computer Vision Systems in the Detection of Heavy Metals in Export Crops with Hyperspectral Imaging and XRF-IoT Sensors distributed across the different quartiles of scientific journals?
- RQ5. Which countries exhibit higher levels of co-occurrence and collaboration in scientific publications related to Computer Vision Systems and their application in the Detection of Heavy Metals in Export Crops with Hyperspectral Imaging and XRF-IoT Sensors?
- RQ6. What central thematic categories are identified in research that integrates Computer Vision Systems with the Detection of Heavy Metals in Export Crops using Hyperspectral Imaging and XRF-IoT Sensors?

The research questions were designed to combine structural field analysis with technological evaluation. Although some RQs include indicators such as performance metrics, journal stratification, and geographic distribution, these are interpreted as proxies of

methodological maturity, research diffusion, and technological consolidation within the domain. RQ6 is supported by science mapping through the Callon strategic diagram, enabling a structural assessment of thematic centrality and density to identify motor, emerging, and peripheral research themes.

These research questions were formulated to ensure that the review integrates not only the technological aspects of heavy-metal detection but also the methodological, computational, and scientific context in which these studies have been developed. By examining how prior work evaluates model performance, which technological ecosystems support current implementations, and how the existing evidence is positioned within the scientific literature, the review gains a broader perspective on reliability, reproducibility, and field maturity. This comprehensive approach strengthens the interpretive depth of the study and supports a more robust understanding of the current state and future potential of hyperspectral imaging and XRF-IoT systems in export-crop monitoring.

3.2. Information Sources and Search Strategies

To extract relevant content, searches were conducted in databases of recognized academic prestige, considered benchmarks in the dissemination of high-impact scientific knowledge. In parallel, customized search equations were designed with the aim of optimizing information retrieval and ensuring a more exhaustive exploration of the following platforms: Web of Science Core Collection, Scopus, IEEE Xplore, EBSCOhost, and Springer Nature Link. The exhaustive search across all databases was finalized on 11 August 2025.

To increase search sensitivity, systematic synonym selection was implemented, which allowed the expansion of the search scope and the retrieval of a larger and more diverse volume of results. This methodological strategy facilitated the identification of relevant publications related to similar topics, even when they did not explicitly employ our initial search variables. In this way, a broader and more representative spectrum of academic evidence was ensured. The synonyms considered are presented in Table 1.

Table 1. Search descriptors and their corresponding synonyms grouped by conceptual category.

Conceptual Group	Descriptor
Computer Vision System	system artificial vision/computer vision/cv/machine visual perception/image analysis/visual pattern recognition/automatic image interpretation/digital image processing/machine vision/automated visual inspection/industrial visual analysis/deep learning/neural networks/convolutional networks/cnn/artificial intelligence/ai/machine learning/ml/transfer learning/image classification/object detection/semantic segmentation
Detection of Heavy Metals in Export Crops with XRF-IoT Sensors	detection/monitoring/identification/analysis + heavy metals/pollutants/toxic metals/metal traces/pollutant scanning/trace control/metal residue control/metal ions/metal toxicity/mineral residues/heavy elements + export crops/export/food agro-export/agro-export/premium crops/agri-food + hyperspectral imaging/hyperspectral vision/multiband sensors/hyperspectral technologies/spectral scanning/spectroscopy/spectral/remote sensing/hyperspectral/spectral signature + xrf sensor/iot-xrf/portable xrf/sensor network/agriculture 4.0/iot/agricultural iot/precision agriculture

To conduct a systematic search, a general search equation was first designed, integrating terms related to Computer Vision/AI, contaminants/metals, and agriculture. Subsequently, customized equations were derived for each source, adjusting syntax and indexing fields (TITLE-ABS-KEY) to maximize retrieval without sacrificing precision.

General equation:

("artificial vision" OR "computer vision" OR "image analysis" OR "machine vision" OR "deep learning" OR "neural networks" OR "convolutional networks" OR "artificial intelligence" OR "machine learning" OR "transfer learning" OR "image classification" OR "object detection" OR "semantic segmentation") AND (("heavy metals" OR pollutants OR "toxic metals" OR "metal traces" OR hyperspectral OR "metal ions" OR "metal toxicity" OR "mineral residues" OR "heavy elements" OR xrf) AND ("agriculture 4.0" OR "agricultural iot" OR "precision agriculture" OR agriculture)).

The complete Boolean query and its database-specific implementations (Scopus, Web of Science, IEEE Xplore, EBSCOhost, and SpringerLink) are provided in the Appendices A–C to ensure full transparency and reproducibility of the search strategy.

3.3. Identified Studies

The identified studies resulted from the rigorous application of the methodological process, which included the selection of academic databases, the systematic incorporation of synonyms, and the design of customized search equations. This strategy enabled the retrieval of a representative set of publications relevant to the research topic, whose distribution is presented in Figure 1.

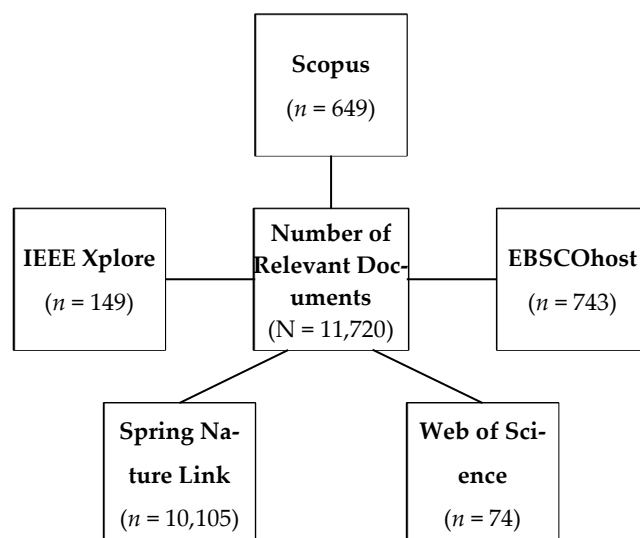


Figure 1. Number of relevant documents.

3.4. Study Selection

The selection of studies was carried out through the application of eight exclusion criteria (EC) aimed at ensuring the relevance, methodological quality, and thematic pertinence of the publications. These criteria allowed for the elimination of papers that did not meet the objectives of the analysis, thereby guaranteeing a systematic refinement consistent with the adopted methodology.

In parallel, a concise set of inclusion criteria was also considered to strengthen transparency in the selection process. Studies were included only if they met the following conditions: (IC1) peer-reviewed academic publication; (IC2) full-text availability; (IC3) publication within the last seven years; (IC4) written in English; (IC5) original and non-duplicated content; (IC6) thematic alignment with heavy-metal detection or the use of hyperspectral imaging, computer vision, or XRF-IoT technologies in agricultural contexts; (IC7) minimum length of ten pages to ensure sufficient methodological depth and reproducibility, as shorter papers -particularly brief conference reports- often lack detailed descriptions of data acquisition, spectral preprocessing, calibration procedures, or model validation required for reliable cross-study synthesis. However, this criterion may intro-

duce a potential exclusion bias against short conference papers and compact sensor-system reports, particularly in IEEE venues, and this limitation is acknowledged when interpreting the findings; and (IC8) abstracts and titles clearly aligned with the research objective.

Additionally, studies were included if they directly addressed heavy-metal detection in agricultural contexts or provided sensing, spectral-processing, or sensor-integration methodologies demonstrably transferable to heavy-metal quantification workflows.

The sequential application of these criteria refined the initial universe of documents and preserved only those studies that were both relevant and methodologically consistent. The process of identification, screening, eligibility assessment, and inclusion was represented through the PRISMA diagram (Preferred Reporting Items for Systematic Reviews and Meta-Analyses), shown in Figure 2, which clearly and transparently illustrates the stages leading to the final selection of the papers incorporated in this review.

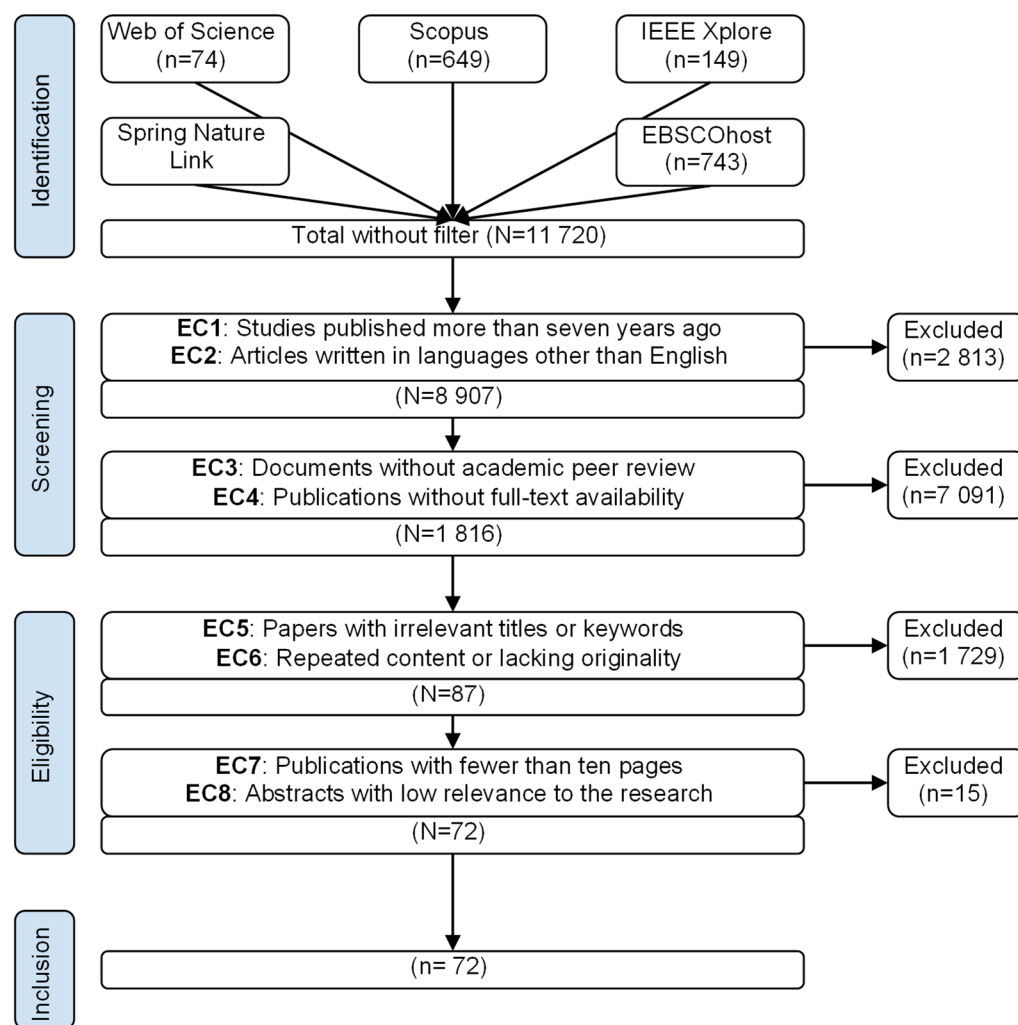


Figure 2. PRISMA flow diagram.

For analytical clarity, the selected studies were categorized into two groups: (i) core studies, directly focused on heavy-metal detection in agricultural or food systems using Computer Vision, hyperspectral imaging, or XRF-IoT technologies; and (ii) methodologically relevant studies, which contribute transferable sensing, spectral processing, or sensor-integration approaches applicable to contaminant detection workflows. This distinction does not alter the inclusion criteria but clarifies the analytical role of each group within the synthesis of findings.

To further strengthen transparency and full alignment with PRISMA 2020, a consolidated matrix summarizing the number of excluded records at each screening stage (EC1–EC8) and for each database has been included in Appendix B (Table A2).

Q4 journal articles and conference papers were included only when they fully met the inclusion criteria, were peer-reviewed, demonstrated direct thematic relevance to heavy-metal detection using HSI or XRF-IoT technologies, and successfully passed the strict quality-assessment protocol applied in this review.

3.5. Quality Assessment

The quality assessment of the included studies was conducted through the application of standardized criteria designed to evaluate the relevance, methodological validity, and scientific rigor of each publication.

- QA1. Is the purpose of the study clearly and explicitly defined?
- QA2. Is the research procedure described with accuracy?
- QA3. Does the paper present a coherent and easily understandable organization?
- QA4. Are the data used in the analysis clearly specified?
- QA5. Do the conclusions align with the stated objectives?
- QA6. Does the study provide an effective response to the formulated problem?
- QA7. Are the results clearly delineated and presented?

To ensure the reliability and validity of the selected papers, a comprehensive evaluation process was applied, based on seven predefined quality criteria. Each criterion was rated as 1 (poor), 2 (acceptable), or 3 (outstanding).

Only papers that reached a global score equal to or higher than 11.5 were deemed suitable for inclusion in the systematic review. This threshold guaranteed a minimum level of methodological rigor and consistency in the evidence incorporated. After the exhaustive review of the 72 initially identified papers, it was verified that all primary studies met this requirement, thereby consolidating the final list of relevant publications. To present the quality assessment in a compact and interpretable manner, the studies were grouped by Type and Quality Score, reporting the frequency and corresponding references for each category. The summary is shown in Table 2, which provides a clear overview of the distribution of quality levels across the 72 studies.

Table 2. Summary of Quality Assessment Results Grouped by Type and Score.

Type	Score	References	Freq.
Journal	15	[9,17,19,26,27,35,48,51–57]	14
Journal	16	[1,11,13,16,25,28,30,31,33,34,49,58–61]	15
Journal	17	[7,8,10,12,15,39,40,62–67]	13
Journal	14	[2,14,20,21,32,42,68–73]	12
Journal	13	[18,74–76]	4
Journal	11	[41,77,78]	3
Journal	12	[44,79]	2
Journal	19	[46,47]	2
Journal	18	[3,43,45,80,81]	5
Journal	10	[82]	1
Conference	15	[83]	1

The database search process initially identified a total of 72 records that met the preliminary inclusion criteria. These records were subjected to a structured quality assessment based on seven specific evaluation questions. As a result of this assessment, four studies [41,77,78,82] were excluded for not achieving the minimum required quality

score (≥ 11.5). Consequently, 68 studies were retained for the final qualitative synthesis and analysis.

3.6. Data Extraction Strategies

The data extraction strategy was implemented using Mendeley Desktop, chosen for its architecture and design that provide an intuitive interface and efficient management of references and documents. These features facilitated the systematic organization of the collected information and optimized the bibliographic analysis process. The outcomes of this process included: (1) structured bibliographic metadata per paper, encompassing title, source, year, country, ISSN, publication type, journal, authors, and institutional affiliation; (2) scientometric indicators, such as indexing database, quartile classification, and citation count; (3) methodological descriptors, including approach, abstract, and keywords; and (4) a unified dataset that ensures transparency, traceability, and consistency across all sources.

Additionally, the research benefited from the use of analytical charts generated with the support of artificial intelligence and the RAj (Research Assistant j) tool developed by Dr. Javier Gamboa Cruzado. This integration of advanced computational tools enhanced data visualization, providing clearer and more precise representations that strengthened the interpretation of results in the analysis and discussion stages.

3.7. Synthesis of Findings

The synthesis prioritizes evidence from core studies directly addressing heavy-metal detection, while methodologically relevant studies are incorporated to contextualize technological developments in sensing and imaging.

For each research question (RQ1–RQ6), the collected information was systematically tabulated and expressed in both quantitative and qualitative terms. This approach enabled more precise comparisons among the different studies analyzed, revealing consistent patterns and emerging trends that have shaped the evolution of research in recent years. Likewise, the combination of both perspectives facilitated a comprehensive interpretation of the findings, thereby strengthening the validity of the final synthesis.

4. Results and Discussion

This section presents the results obtained from the systematic review and analyzes them in relation to previous studies and the objectives defined for this research. The comparison between the collected evidence and the initial objectives enables the identification of advances, limitations, and knowledge gaps in the subject under study. The discussion prioritizes an integrative interpretation rather than isolated metric comparisons, aiming to reveal technological convergence, methodological trends, and research gaps across the analyzed studies.

4.1. General Description of the Studies

Figure 3 illustrates the geographical distribution of the published papers through a world map and a bar chart. This visualization allows the identification of both the spatial concentration of scientific production and the relative contribution of the countries most active in the research topic.

The results show that China clearly leads production with 24 papers, followed by the United States (9) and India (6), reflecting the consolidation of these countries as research hubs in technologies applied to food security and computer vision. Europe presents a more fragmented contribution, with Italy, France, Spain, and Hungary as prominent representatives, while in the Americas, contributions are concentrated in the United States and Brazil. Likewise, Middle Eastern and North African countries such as Saudi Arabia

and Egypt are emerging with a relevant role in recent production. The distribution reveals a pattern of global asymmetry, in which most scientific knowledge is concentrated in specific regions, leaving significant gaps in Sub-Saharan Africa and much of Latin America, which limits geographic diversity in the generation of evidence.

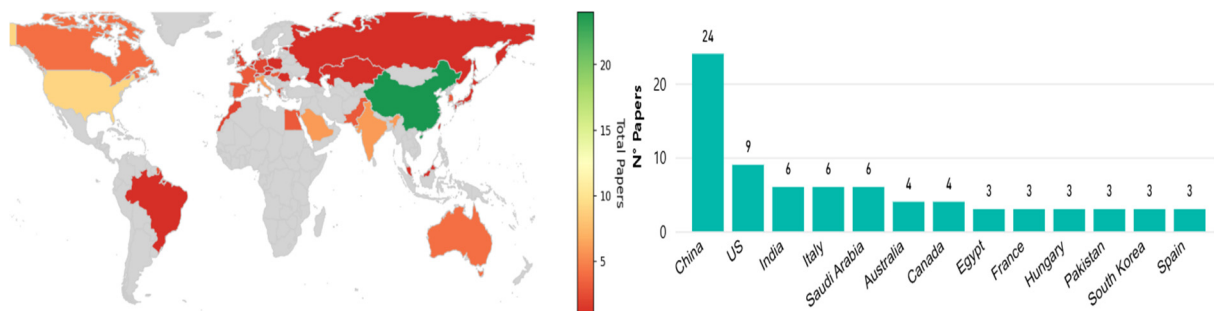


Figure 3. Distribution of papers by country.

International evidence converges on the fact that scientific production is concentrated in a few hubs: M. R. Sabour, P. Sakhaie, and F. Sharifian [84] underline China and India as the main contributors; in the same vein, Juan Botero-Valencia and collaborators [24] confirm this leadership through citation metrics and publication trends; Nayeli Montalvo-Romero and colleagues [85] specify that China (36) and India (12) lead the output, followed by the United States, Italy, and Pakistan; complementarily, Josef Souza Silva and other authors [86] show that China, India, and the United States account for more than 20 publications, with scattered contributions (1–5) across the rest of the countries; and Xionghai Chen and his team [87] document China’s exponential growth ($\geq 35\%$ between 2018 and 2023) compared to the low and stable participation of Belgium, Canada, and France. Finally, R. Bertoglio and their research group [22] detail high absolute volumes in the United States (653) and China (621) in digital agriculture, reflecting technological maturity and regional concentration of capacities.

This scenario suggests the need to foster the transfer of methodologies to other business sectors such as agribusiness, mining, and healthcare, where contaminant monitoring is equally critical. The limited participation of certain regions opens opportunities to expand research into new geographical areas, adapting models to diverse local conditions. Furthermore, the periodic replication of these studies would allow the identification of the evolution of trends over time, strengthening predictive capacity in technological development. Finally, international cooperation between leading countries and those with lower production is projected as a key factor to democratize access to knowledge and reduce gaps in applied research.

Figure 4 presents the temporal evolution of scientific production, represented through a pie chart and a Pareto chart. This dual visualization makes it possible to identify both the relative proportion of papers per year and the cumulative concentration of production in specific periods.

The results show sustained growth in recent years, with a notable increase starting in 2021 and a peak in 2024 with 21 papers (30.88%). The year 2021 stands out as the second most productive period with 13 publications (19.12%), followed by 2023 with 10 (14.71%). Production in 2025, although partial, already records 11 papers (16.18%), which anticipates a continuation of the upward trend. In contrast, the years prior to 2020 show very limited production, with only 2 papers in 2019 and 4 in 2020, reflecting a field in recent consolidation. The Pareto chart confirms that 2024, 2021, and 2025 together account for more than 65% of total production, evidencing a rapid maturation of the topic. Overall,

the data suggest that research on computer vision and IoT for heavy metal detection has strongly emerged over the past half decade, consolidating as a topic of increasing relevance.

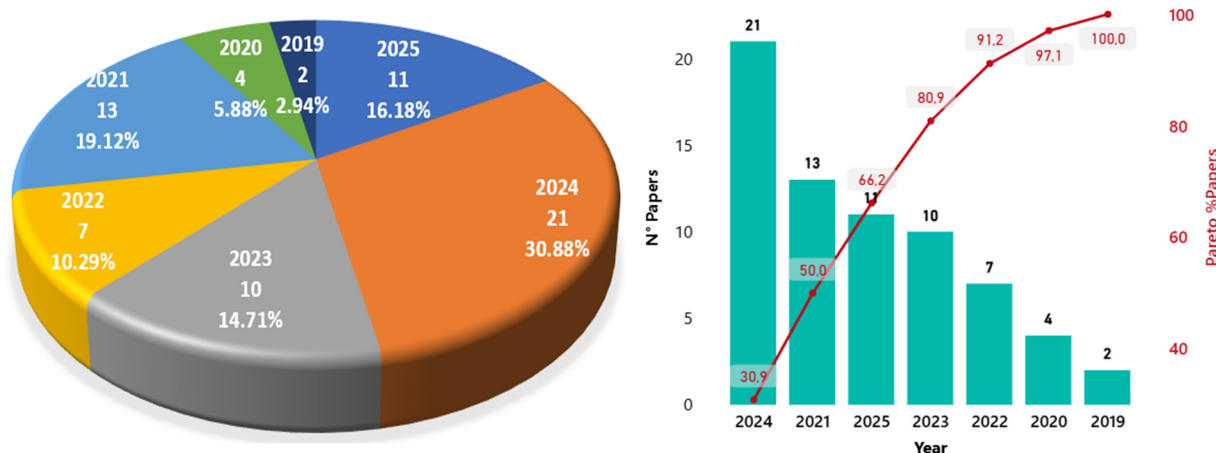


Figure 4. Distribution of papers by year.

These findings suggest that the field is expanding and holds potential for application in other business sectors such as agribusiness, mining, and environmental health. The temporal concentration opens the possibility of extending research to underrepresented geographical regions, while the continuity of studies will allow evaluation of the sustainability of the trend across different time horizons. Finally, this growth pattern highlights the importance of international cooperation to expand the applicability and transfer of these technologies to diverse contexts.

Figure 5 and Table 3 show the relationship between scientific production classified by quartiles and year, and its impact measured in citations and bibliometric indicators, which makes it possible to assess not only the quantity of publications but also the quality and relevance of the studies over time.

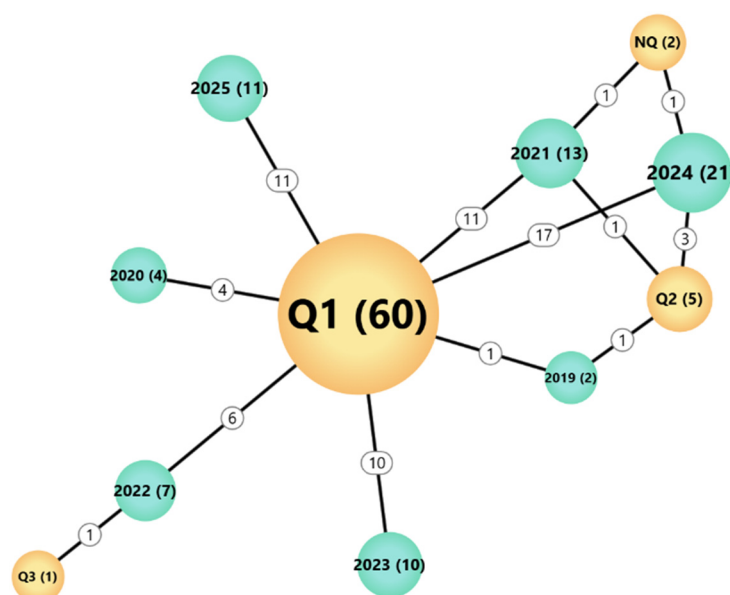


Figure 5. Papers by year and quartile. Note: Yellow nodes represent journal quartiles; green nodes represent publication years; numbers inside the nodes indicate the number of papers.

Table 3. Impact of publications by year.

Year	N° Papers	Papers (%)	N° Citations	Citations (%)	H-Index	H-Index (%)	Citations/Paper
2024	21	30.9%	99	6.6%	3411	25.8%	5
2021	13	19.1%	721	48.0%	2472	18.7%	55
2025	11	16.2%	7	0.5%	2227	16.9%	1
2023	10	14.7%	264	17.6%	2062	15.6%	26
2022	7	10.3%	156	10.4%	1718	13.0%	22
2020	4	5.9%	201	13.4%	902	6.8%	50
2019	2	2.9%	55	3.7%	415	3.1%	28
Total	68	100.0%	1503	100.0%	13,207	100.0%	22

The results indicate that the largest production is concentrated in Q1, with 60 papers distributed mainly between 2021 and 2024, confirming the authors' preference for publishing in higher-impact journals. The year 2024 recorded the highest number of papers (21), although with a relatively low citation impact (6.6%), while 2021, with 13 publications, accounted for 48% of total citations, reflecting its high academic visibility. In contrast, 2025 shows a volume of 11 papers but still with little citation impact due to its recent publication. The years 2019 and 2020 show more limited production, though with a high citation index per paper, evidencing pioneering contributions in the field. The H-index analysis confirms that the periods 2021 and 2024 were decisive in consolidating the field, jointly contributing nearly half of the accumulated impact. Overall, the data reveal a dynamic of sustained growth in production and differentiated impact, where the quality and timeliness of papers determine their influence in the scientific community.

V. Yele, R. Sedamkar, and S. Alegavi [9] report sustained growth of Deep Learning in remote sensing between 2014 and 2023. Complementarily, U. Khan and collaborators [76] conducted a rigorous selection process in which, after filtering 2200 papers, they identified 44 primary studies; their findings show a surge in 2020 with 16 publications, reflecting the growing interest in applying Deep Learning and hyperspectral technologies in agriculture. On the other hand, S. Swaraj and S. Aparna [18] point to an upturn beginning in 2022, driven by technical advances and emerging trends. Likewise, M. F. Guerri and colleagues [6] describe a steady increase since 2017, with a marked rise in 2021–2022, when publications exceeded 40 papers. In turn, R. Bertoglio and their team [22] identify exponential growth since 2015, reaching a peak in 2019 that reflects the maturation of the technology and its expansion into broader applications. Finally, T. Andrade-Mogollón and co-authors [23] highlight that in 2024, 28 publications were recorded, mainly in Scopus, consolidating the current expansion phase in the field.

These findings suggest that publishing in Q1 journals not only increases academic visibility but also constitutes a strategic vehicle for transferring knowledge to business sectors such as agribusiness, mining, and health. The variation in impact by year reinforces the need to replicate research in other geographical regions and across different time horizons to diversify the available evidence. Moreover, the asymmetric citation behavior underscores the importance of fostering long-term studies that transcend temporary trends and generate sustained impact. Finally, the analysis highlights the opportunity to establish international collaboration networks to strengthen quality and broaden the applicability of these technologies in global contexts.

To improve transparency regarding the technological composition of the evidence base, a quantitative characterization of the 68 primary studies was conducted. Table 4 summarizes the distribution of studies according to their focus on heavy metal detection, involvement in agricultural contexts, and the sensing technologies employed.

Table 4. Technological characterization of primary studies.

Category	N° Papers
Direct heavy metal detection	18
Export crops/agricultural context	58
Use of Hyperspectral Imaging (HSI)	42
Use of XRF/XRF-IoT	1
Combination HSI + XRF	1

To address the sensor-level evidence requested by the reviewer, the studies that directly estimate heavy metals using HSI/Vis–NIR spectroscopy, together with the XRF-related evidence identified within the primary set, are systematically characterized in Appendix C Table A3, including platform, spectral range, spatial resolution, target metals, and reported performance.

As observed in the results, the lower proportion of studies focused on direct heavy metal detection reflects the technical complexity and costs associated with chemical validation. According to the observed distribution, the dominance of agricultural contexts and the widespread use of HSI can be explained by the consolidation of precision agriculture, greater sensor availability, and the existence of standardized datasets that facilitate adoption. In contrast, the limited presence of XRF/XRF-IoT and its scarce integration with HSI suggest operational and technological barriers that still constrain the development of robust multimodal approaches.

Figure 6 and Table 5 show the frequency and relevance of the keywords used in the papers analyzed, represented in a word cloud and in their distribution by quartile. This approach makes it possible to identify the dominant conceptual axes and the quality of the publications according to the impact level of the journals.

**Figure 6.** Word cloud of keywords from the analyzed papers.

The results show that the most frequent words are machine learning (ML) with 20 occurrences, followed by hyperspectral imaging (13), precision agriculture (12), and deep learning (10), confirming the centrality of artificial intelligence–based approaches in recent research. Likewise, terms such as remote sensing (7) and agriculture (6) reinforce the connection of these technologies with direct applications in the agricultural sector. Specific concepts such as CNN, random forest, and soil contamination appear less fre-

quently, although they represent methodological and applied lines of growing interest. The distribution by quartile reveals that the vast majority of these publications (344 out of 382) are concentrated in Q1 journals, reflecting the high quality and impact of the field. Furthermore, the presence of terms associated with environmental issues, such as heavy metals and soil contamination, highlights the importance of this research in sustainability and food security.

Table 5. Distribution of keywords by quartile.

Keyword	NQ	Q1	Q2	Q3	Total
ml	0	18	1	1	20
hyperspectral imaging	0	12	1	0	13
precision agriculture	0	12	0	0	12
deep learning	0	9	1	0	10
remote sensing	0	7	0	0	7
agriculture	0	6	0	0	6
cnn	0	3	0	0	3
convolutional neural networks	0	2	1	0	3
convolutional neural networks (cnn)	0	2	1	0	3
heavy metals	0	3	0	0	3
random forest	0	3	0	0	3
soil contamination	0	3	0	0	3
ai	0	2	0	0	2
chemometrics	0	2	0	0	2
classification	1	1	0	0	2
crop classification	0	2	0	0	2
...	...	...	...	...	...
Total	10	344	24	4	382

In this study, the predominant terms are machine learning, artificial intelligence, deep learning, and internet of things, confirming the centrality of these technologies in precision agriculture. Alvarado and collaborators [88] developed a word cloud based on the most frequent terms in Machine Learning-based solutions. The most prominent terms -machine learning, artificial intelligence, deep learning, and internet of things- highlight their close interconnection, while concepts such as neural networks, predictive maintenance, and natural language processing illustrate the diversity of technological applications. Similarly, José Castro-Espichan and colleagues [89] visually display the most frequent terms without the need to exclude stop words, since the analysis focuses only on titles; among the most representative words are impaired, vision, augmented, and reality, which reflect the key concepts of the study and coincide with the terms used in the search equations. In the same vein, Aléxis Cárdenas-Quispe and collaborators [90], who also identify machine and learning as dominant concepts, although in their case linked to threat detection in Android devices, evidencing that machine learning constitutes a cross-cutting axis in different areas of research beyond the agricultural field.

The predominance of keywords linked to artificial intelligence and precision agriculture suggests broad potential for transfer to business sectors such as mining, health, and environmental management. The concentration in Q1 journals reinforces the legitimacy of the field, opening opportunities to expand these studies to other geographical regions that still show limited production. The evolution of these keywords over time will make it possible to detect emerging trends and guide future research agendas. Finally, the integration of environmental and technological terms highlights the possibility of developing interdisciplinary solutions with applications in different contexts and production scales.

Table 6 presents the 15 institutions with the greatest impact on scientific production, considering the number of papers published, citations received, accumulated H-Index, and average citations per paper, allowing the relevance and visibility of each academic entity in the field to be evaluated.

Table 6. Institutions with the greatest impact on publications.

Affiliation	N° Papers	N° Citations	H-Index	Citations/Paper
Ghent University	2	81	614	41
Guizhou Medical University	2	6	798	3
King Khalid University	2	8	429	4
Kunming University of Science and Technology	2	12	262	6
Mohammed VI Polytechnic University	2	36	434	18
RMIT University	2	50	371	25
Tsinghua University	2	33	443	17
Xinjiang University	2	56	338	28
Yeungnam University	2	1	407	1
Aarhus University	1	1	274	1
Alexandria University	1	6	59	6
AlMaarefa University	1		139	
Anhui University	1	5	0	5
Beijing Normal University	1	5	274	5
California State University Fresno	1	5	98	5
Central China Normal University	1	44	217	44
...	...	...	...	...
Total	163	3302	31,085	20

The data show that Ghent University leads in citations per paper with an average of 41 (81 citations across 2 papers, H-Index 614), while Central China Normal University also achieves remarkable performance with 44 citations per paper (44 citations in 1 paper, H-Index 217). Xinjiang University stands out with 56 citations in 2 papers (28 per paper, H-Index 338), followed by RMIT University with 50 citations in 2 papers (25 per paper, H-Index 371) and Mohammed VI Polytechnic University with 36 citations in 2 publications (18 per paper, H-Index 434). In contrast, institutions such as Guizhou Medical University (6 citations in 2 papers, 3 per paper, H-Index 798) and Yeungnam University (1 citation in 2 papers, 1 per paper, H-Index 407) show relatively lower impact despite high accumulated H-Index values. Universities with only one paper, such as Aarhus University (1 citation, H-Index 274) and Beijing Normal University (5 citations, H-Index 274), reflect punctual contributions, while King Khalid University (8 citations in 2 papers, 4 per paper, H-Index 429) and Kunming University of Science and Technology (12 citations in 2 papers, 6 per paper, H-Index 262) demonstrate active production albeit with moderate citation counts. Altogether, the 15 institutions analyzed account for 163 papers and 3302 citations, with a general average of 20 citations per paper and an accumulated H-Index of 31,085, confirming the consolidation of the field in strategic research hubs with institution-specific differentiated impact.

These findings suggest the importance of fostering collaborations between institutions leading in citations per paper and those with high H-Index, integrating quality and volume to strengthen research agendas. The concentration of leadership in Chinese and European universities highlights the need to expand production to underrepresented regions such as Latin America and Sub-Saharan Africa in order to diversify contexts and validate results. Replicating this analysis in different periods will make it possible to track the evolution of institutional impact and detect emerging research hubs. Finally, the lessons drawn from these dynamics can be transferred to business sectors such as agribusiness, health, and

mining, where international cooperation and the development of knowledge networks are key to boosting innovation and technological transfer.

Table 7 synthesizes the key findings of the 68 studies analyzed, grouped by pre-dominant methodological category. This organization makes it possible to identify recurring patterns in methods, datasets, evaluation metrics, and limitations, facilitating cross-comparison and the extraction of general lessons. Additionally, this synthesis explicitly links sensing architectures, reported performance metrics, and operational deployment scenarios -including laboratory, field, UAV-based, and multimodal environments- thereby reinforcing the technological core of the review. Rather than serving as a benchmarking exercise, this synthesis identifies cross-study patterns in sensing architectures, dataset dependencies, performance behavior, and operational constraints, enabling a structural interpretation of technological maturity within the field. In addition, only the most relevant studies for the aim of this work were included, prioritizing: (i) explicit alignment with the detection of heavy metals in export crops through hyperspectral imaging and/or XRF-IoT sensors; (ii) methodological robustness and reporting of comparable metrics (e.g., RMSE/MAE, OA, R^2 , AUC), along with reproducible validation protocols (cross-validation, external validation, or field trials) and sufficient data description; (iii) dataset traceability; and (iv) consistency of results. Studies with non-comparable metrics were excluded, ensuring traceability through the relation of identifiers by category.

Method category: The taxonomy separates cores aligned with the objective (DL–HSI and ML for heavy metals), connected fronts (spectral engineering, risk mapping, Unmanned Aerial Vehicle (UAV)/multimodal), and peripheral segments (environmental/other, out of scope). This configuration reveals a structurally uneven technological landscape in which mature HSI classification pipelines coexist with still incipient efforts in operational heavy-metal quantification and monitoring.

Methods used: DL architectures for HSI predominate (3D-UNet, CNN-LSTM, attention, self-supervised), while in metals, classical ML methods (RF/PLSR/SVR, ensembles, fractional derivatives) with simple fusions remain prevalent. This methodological split reveals that advances in spectral–spatial representation have outpaced the development of scalable and deployment-ready regression pipelines for contaminant quantification.

Datasets: There is strong dependence on benchmarks (AVIRIS, Pavia, WHU-Hi, UAVINE) and soil databases (LUCAS, Henan, Suzi river, Lianzhou), with few fully characterized operational UAV scenarios. This data concentration toward laboratory or well-instrumented environments limits cross-site generalization and constrains transferability to heterogeneous agroecosystems and industrial monitoring contexts.

Performance: Maximum values (OA $\approx$ 99–100%, AUC $\approx$ 0.96, R^2 up to 0.96) are typically achieved under controlled conditions and classification tasks. However, field and soil applications exhibit larger error dispersion (higher RMSE, OA degradation, and lack of standardized metrics in operational deployments), revealing persistent heterogeneity and limited normalization across studies.

Limitations: Recurrent constraints include scarcity of labeled data, limited generalization, spectral similarity between classes, calibration and environmental variability, spatial representativeness issues, and spectral mixing effects. Additionally, only a minority of studies report computational burden, scalability, or real-time feasibility, highlighting a persistent gap between laboratory-level accuracy and industrially deployable heavy-metal monitoring systems.

Table 7. Synthesis of findings by methodological category.

Method Category	Methods Used	Datasets	Performance	Limitations	Refs.	Qty. (%)
DL–HSI (classification/segmentation)	CNN/ResNet/3D-UNet, CNN-LSTM, self-supervised contrastive, attention; DeepLabV3+	AVIRIS (Cuprite, Salinas, Indian Pines), Pavia, WHU-Hi (LongKou, HanChuan), UAVINE, field/lab images	OA/Accuracy 69.7–99.8%; Kappa up to 0.89; Precision up to 100% (field tends to underperform lab)	Scarce labels; limited generalization beyond datasets; spectrally similar classes; cost/hardware dependence	[1–3,7,16,28,30,31,40,41,53,80]	12 (20.34)
ML for heavy metals (regression/classification)	RF/PLSR/SVR/Ensembles (AdaBoost, GAMLSS-RF), fractional derivatives; DA-SVM/DNN	Soils: LUCAS, Henan, Suzi river, Lianzhou; multi-site soils/sediments	$R^2 = 0.46\text{--}0.96$; RMSE = 0.26–49.68 mg/kg; Accuracy up to 96%	Spatial representativeness; lower accuracy for trace elements; spectral noise; sampling bias	[46,52,56,57,68,75,78]	7 (11.86)
Soil/vegetation properties (non-metals)	SVR, PLS/PLSR, LWR, RF, AutoML; water/strain/soil indices; spectral regression	HSI/NIR-SWIR UAV/aerial; in situ leaf reflectance; Kuusiku (EE); multiple fields	$R^2 = 0.38\text{--}0.99$; RMSE = 0.11–0.36 (various targets); Accuracy up to 96%	Environmental/temporal variability; calibration; preprocessing needs; indicator dependence	[11,12,34,42,44,47,54,62,63]	9 (15.25)
Spectral engineering/indices and unmixing	NFINDR, index design (HBSI, NSI); simple fusion (RF/SVM)	AVIRIS-NG, Sentinel-2; multiple sites	Accuracy 93.5–98.6%; Kappa 87.8–97.4%; Unmixing CS < 5%; NSI responds in 10/13 datasets	High dimensionality; scarcity of public HSI datasets; over-segmentation; index assumptions	[43,45,71]	3 (5.08)
Environmental risk mapping (pesticides/contaminants)	RF, PMF + RF, Monte Carlo; meta-analysis; ensembles	2307 sites (309 pesticides), Wenling; multi-study meta data	AUC 0.96; Accuracy up to 95.5%; R^2 up to 0.86; meta: –66% concentration in organisms	Uneven monitoring; mixture effects; sampling biases; design heterogeneity	[14,15,81–83]	5 (8.47)
Operational remote sensing (UAV/multimodal)	UAV HSI; multimodal integration + AI	Field UAV HSI; operational scenarios	Reported accuracy gains; improved real-time efficiency (no common metric)	Calibration; robust training datasets; environmental variability	[74]	1 (1.69)
Environmental/other (not directly HSI–metals)	ML for water quality, bioaccumulation, adsorbents; RSM/ANN; LULC mapping	Rivers (Nakdong), water/plant samples, spectroradiometer; land-use data	Heterogeneous (e.g., R^2 up to 0.999; high BAFs)	Limited transfer to target objective; domain differences; comparability issues	[13,58,59,64,69,79]	6 (10.17)
Out of scope (NLP/Chatbots/Dialog)	LSTM/Seq2Seq with attention; decision frameworks (CoCoSo/AHP) in telco	Dialogue datasets; non-agricultural cases	N/A for the paper’s objective	Do not evaluate Acquired Vegetation Samples (AVS)/HSI or metals in crops; domain incompatibility	[10,19–21,27,32,39,51,55,60,61,65–67,72,77]	16 (27.12)

From an integrated review perspective, the reported performance figures across studies should not be interpreted as directly comparable benchmarks, since they are derived from heterogeneous datasets, experimental conditions, and validation protocols. Furthermore, only a limited subset of works reports computational burden, model scalability, or real-time feasibility—critical factors for industrial adoption in precision agriculture. This evidence indicates that, despite strong progress in spectral–spatial modeling accuracy, the field still exhibits a notable gap between laboratory-level performance and deployment-ready heavy-metal monitoring systems in export-crop environments.

The results recommend standardizing core metrics per task (e.g., OA/F1 for classification; RMSE/ R^2 for regression) and consistently reporting field scenarios alongside laboratory experiments to quantify transfer gaps. It is necessary to expand and diversify datasets (multi-site, multi-temporal, multi-crop) with reproducible calibration protocols and semi-/self-supervised annotation strategies to reduce label dependence. The combination of DL–HSI with explainable ML for metals (e.g., RF/PLSR with SHAP) can balance performance and interpretability. External validation and normalization of HIS + XRF–IoT fusions should be prioritized for robust quantification. The scarcity of operational UAV/multimodal studies suggests investing in real-time pipelines and domain adaptation. Finally, finer thematic alignment filters will improve the usefulness of synthesized evidence for food safety and agro-export competitiveness policies.

4.2. Answers to the Research Questions

This section addresses the research questions by presenting the main findings, their discussion, and the implications for future research. The results derive from a rigorous bibliographic search aimed at identifying journal and conference papers on the application of computer vision systems for detecting heavy metals in export crops using hyperspectral imaging and XRF-IoT sensors. The search was conducted across well-established academic databases and, following an exhaustive screening and quality assessment to ensure source relevance, only the most significant studies were retained. This filtering process resulted in a final corpus of 68 papers, whose findings are analyzed in relation to the research questions defined in this study.

Beyond the bibliometric structuring, the synthesis explicitly integrates system typologies, sensing configurations, performance indicators, methodological limitations, and deployment contexts of HSI- and XRF-based approaches applied to heavy-metal detection in agricultural and soil environments.

RQ1: *What performance indicators are used to measure the effectiveness of computer vision systems?*

Although the metrics identified are widely known in computer vision, analyzing them within this specific domain reveals how models for heavy-metal detection are actually being evaluated. This allows us to identify methodological patterns—such as the strong dependence on error-based metrics and the limited use of operational indicators—that directly affect comparability, maturity, and real-world applicability of HSI and XRF-IoT systems.

Table 8 and Figure 7 show the distribution of criteria used to evaluate the effectiveness of computer vision systems in detecting heavy metals in export crops. These metrics allow for the comparison of accuracy, efficiency, and reliability of the models, establishing a solid foundation for the validity of the results obtained.

Table 8. Criteria for measuring computer vision systems.

Evaluation Criterion	Reference	Qty. (%)
Root Mean Square Error (RMSE)	[10–12,20,32,34,39,42,45,46,48,54,56,59,64,65,74,77,81]	19 (67.9)
Mean Absolute Error (MAE)	[3,20,39,58,59,82]	6 (21.4)
Inference Time	[3,16]	2 (7.1)
ROC Curve	[53]	1 (3.6)

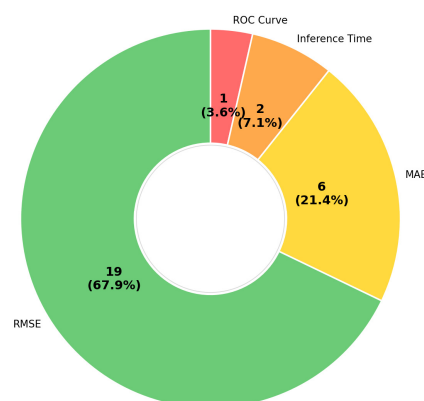


Figure 7. Distribution of evaluation metrics.

The results indicate a predominance of the Root Mean Square Error (RMSE), reported in 19 studies (67.9%), reflecting its acceptance as a standard metric for quantifying deviations and assessing accuracy in continuous predictions. The Mean Absolute Error (MAE) appears in 6 publications (21.4%), serving as a complementary measure to capture absolute errors without considering their direction, thus adding robustness to the evaluation. To a lesser extent, Inference Time, reported in 2 studies (7.1%), highlights the relevance of computational efficiency in real-time application contexts. Finally, the ROC Curve, used in a single study (3.6%), reveals a limited interest in classification-related metrics within this specific domain. This distribution demonstrates a clear orientation toward error metrics in regression models, relegating classification indicators to a marginal role. Overall, the findings show a preference for traditional quantitative measures, although with emerging attention to operational parameters such as inference time.

Beyond the descriptive distribution of metrics, the evidence shows a structural limitation in current evaluation practices. The strong reliance on RMSE and MAE indicates that most studies assess performance only through numerical error, without considering robustness, reproducibility, or stability under real-world variability, key aspects in XRF-based detection. The absence of probabilistic or operational indicators (e.g., AUC, precision–recall, false-negative rates, or inference latency) also restricts the ability to determine real deployment feasibility. Thus, current evaluation frameworks remain incomplete: they quantify error but do not capture practical reliability or field readiness. A more mature framework should integrate robustness metrics, uncertainty estimation, and operational criteria to ensure meaningful comparability and real-world applicability of XRF-IoT detection systems.

V. Yele, R. Sedamkar, and S. Alegavi [9] point out that the accuracy metric is the most widely used to calculate the performance level of the techniques applied. In line with this, A. Khan and collaborators [91] identify both accuracy and R^2 as the most frequent indicators for measuring the effectiveness of their models. Likewise, S. Swaraj and S. Aparna [18] highlight the prevalence of accuracy, although the F1-score also appears. On the other hand, B. Ram and colleagues [92] emphasize that accuracy was the only metric applied in their research, reflecting a restricted approach to performance evaluation. Finally, K. Yu and co-authors [5] confirm that in the field of computer vision, accuracy and R^2 predominate as reference metrics. Taken together, these studies underline that literature tends to privilege classification indicators, whereas in the corpus analyzed a greater use of regression metrics is observed, thus evidencing a relevant methodological difference between contexts.

The strong reliance on error metrics such as RMSE and MAE suggests that these tools can be successfully extrapolated to other sectors that rely heavily on numerical predictions, such as energy, healthcare, or environmental management. The growing incorporation of inference time opens opportunities in business applications that require fast, real-time responses, such as export-oriented agribusiness or industrial monitoring. Likewise, the limited use of classification metrics such as the ROC Curve indicates potential for expanding their application to other geographic and temporal scenarios where contaminant classification is a priority. Finally, these results highlight the need to combine error, efficiency, and classification metrics in future research in order to achieve more comprehensive and comparable evaluations across different productive contexts.

RQ2: *Which programming languages are most frequently used in the development of computer vision systems?*

Even though the predominance of Python and the presence of MATLAB or Scala are well-known trends, examining them in this context helps uncover technological dependencies that influence reproducibility and scalability. The findings show how the choice

of programming language shapes the practical feasibility of deploying HSI and XRF-IoT detection systems in real agricultural environments.

Table 9 and Figure 8 present the distribution of programming languages employed in the reviewed studies, allowing for the identification of trends and technological preferences in the development of computer vision systems applied to the detection of heavy metals in export crops. These results provide a clear overview of the tools most widely adopted by the research community.

Table 9. Programming languages used.

Programming Language	Reference	Quantity (%)
Python	[1,15,31,40–42,44,51,54,58,62,67,75,80,82]	15 (41.6)
C++	[80]	1 (2.8)
MATLAB	[12,13,28,31,43,44,52,56,59]	9 (25)
Java	[83]	1 (2.8)
Scala	[10,15,25,44,45,53,54,60,74]	9 (25)
Swift	[42]	1 (2.8)

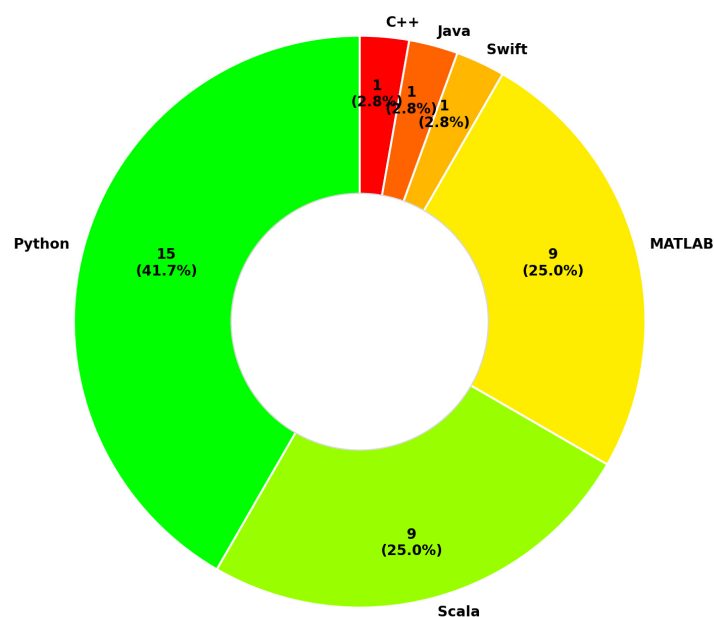


Figure 8. Distribution of programming languages.

The findings show the predominance of Python, with 15 references (41.6%), confirming its role as the standard language in computer vision due to its wide availability of specialized libraries and frameworks. MATLAB and Scala appear with equal frequency, each in 9 publications (25%), underscoring their relevance both in academic environments for modeling and simulation (MATLAB) and in applications oriented toward big data and distributed processing (Scala). The presence of C++, Java, and Swift is marginal, with only one reference each (2.8%), suggesting their use in specific contexts such as performance optimization (C++), cross-platform applications (Java), or mobile development (Swift). This distribution reflects a diversified ecosystem, but one clearly dominated by Python as

the reference language, complemented by options with more focused applications. Taken together, the results highlight a convergence toward flexible and scalable environments that integrate both academic research and practical deployment.

Beyond the descriptive distribution of languages, the evidence shows a strong technological dependence that affects reproducibility and field deployment. Python's dominance concentrates development in a single ecosystem, facilitating experimentation but limiting diversification for real agricultural environments. The marginal presence of C++, Java, and Swift reveals little focus on real-time or embedded implementations—crucial for HSI and XRF-IoT systems. Likewise, the parallel use of MATLAB and Scala reflects a split between academic prototyping and big-data processing. Overall, current choices support early experimentation but fall short in portability, optimization, and scalability for real-world applications.

S. Swaraj and S. Aparna [18] identify Python and MATLAB as the most widely used languages, surpassing alternatives such as Raspberry or Arduino. Complementarily, T. Andrade-Mogollón and colleagues [23] emphasize that development ecosystems based on Python and C/C++ are the most influential in implementing IoT solutions aimed at precision agriculture. In the same line, Sahu and collaborators [93] confirm the use of MATLAB and Python, while J. Akbar and co-authors [94] underline the usefulness of Python for building architectures across different models. In turn, Khan, Laghari, and Awan [91] point out that Python and Java predominate in the field of machine learning. These findings align with studies applied to chatbots, where Albites-Tapia and their research group [95] highlight the dominance of Python in medical diagnostics due to its support for artificial intelligence and backend processing, and Gamboa-Cruzado and his team [96] confirm its prevalence (30.2%) in customer service chatbots. Taken together, these findings reinforce what is observed in this study, where Python leads with 41.6% of the references, supported by a robust ecosystem of libraries (OpenCV, TensorFlow, Scikit-learn), while MATLAB and Scala maintain a significant presence (25% each), and languages such as C++, Java, and Swift appear only marginally.

The leadership of Python implies that its advantages can be transferred to other sectors such as healthcare, logistics, or manufacturing, where computer vision is becoming increasingly consolidated. The use of MATLAB and Scala opens the possibility of strengthening projects in geographic regions where universities and industries already rely on these platforms. The low presence of C++, Java, and Swift highlights a space of opportunity for their future integration, particularly in real-time, mobile, or high-performance applications. Finally, the temporal evolution of these preferences will allow researchers to anticipate changes in the technological ecosystem and guide the training of human capital toward languages with greater projection in both research and the business sector.

RQ3: *What are the main conceptual trends in research on computer vision systems and their impact on the detection of heavy metals in export crops with hyperspectral imaging and XRF-IoT sensors?*

Table 10 presents the most frequent bigrams in the reviewed studies, enabling the identification of conceptual trends that have guided recent research on computer vision systems applied to the detection of heavy metals in export crops. This visualization contributes to recognizing the most dynamic areas and the dominant methodological connections in the field.

The results reveal that machine learning and remote sensing stand out as the most representative bigrams, consolidating themselves as conceptual pillars of research. The recurrent presence of terms such as deep learning, neural networks, and random forest confirms the centrality of artificial intelligence-based approaches in the classification and prediction of contaminants. Concepts such as feature extraction and hyperspectral data

highlight the importance of advanced spectral data processing techniques to improve detection accuracy. Likewise, bigrams associated with the application context, such as study area, organic matter, or land cover, reflect the effort to link methodologies with specific environmental and agricultural conditions. Taken together, this distribution suggests a transition from general machine learning approaches toward more specialized methodologies that integrate spectral analysis with robust AI models, strengthening early and non-invasive detection of heavy metals.

Table 10. Number of concepts (bigrams) per year.

Bigrama	2019	2020	2021	2022	2023	2024	2025	Total
machine learning	0	1	6	3	6	10	8	34
remote sensing	0	3	6	2	7	11	5	34
deep learning	0	1	4	1	3	8	3	20
neural networks	0	1	3	0	4	7	4	19
neural network	2	0	4	1	3	7	1	18
hyperspectral data	0	1	3	1	3	6	2	16
random forest	0	0	2	2	2	6	4	16
feature extraction	0	1	1	0	3	6	4	15
study area	0	0	4	3	2	2	3	14
hyperspectral images	0	0	2	2	3	2	3	12
organic matter	0	0	6	2	0	3	1	12
spectral bands	0	1	1	1	2	4	3	12
support vector	0	2	1	0	3	3	3	12
convolutional neural	0	0	3	0	0	6	2	11
heavy metal	0	0	4	2	1	4	0	11
...	...	...	...	...	...	...	...	...
Total	87	184	600	323	457	986	514	3151

Beyond the descriptive distribution of bigrams, the evidence shows that research in this domain is consolidating around a stable set of methodological pillars. The dominance of machine learning, remote sensing, and deep learning confirms a clear shift toward AI-driven modeling, while the recurrent presence of hyperspectral data, feature extraction, and spectral bands highlights the centrality of advanced spectral processing for detecting contaminants in crops. However, the trends also reveal a conceptual limitation: despite the rapid evolution of XRF-IoT technologies, most studies still rely on traditional learning pipelines and do not incorporate more recent architectures or multimodal strategies that could better capture spectral–chemical interactions. This suggests that conceptual development, although growing, remains anchored in conventional frameworks and has not yet transitioned toward the next generation of sensing–AI integration. Strengthening this line of research requires broadening the methodological spectrum, integrating more robust spectral models, and expanding the environmental contexts where these techniques are validated.

Syeda I. Hassan and colleagues [4] report that the most widely used technology is multispectral, a direct precursor of hyperspectral technology, confirming the relevance of remote sensing in this field. On the other hand, Guerri and co-authors [6] highlight the growing application of CNNs and GANs, deep learning architectures that have gained prominence in computer vision. In concordance with our findings, R. Bertoglio and their research group [22] emphasize an exponential increase in deep learning in recent years, reflecting the consolidation of emerging technologies with greater processing capacity. These results correspond to what is observed in this study, where the bigrams machine learning, remote sensing, and deep learning emerge as the most representative, evidencing a clear convergence toward methodologies based on artificial intelligence and advanced sensing. However, Abderahman Rejeb and collaborators [97] present an extensive list

of bigrams, with soil temperature and air temperature standing out, and subsequently illustrate their distribution and frequency, providing a detailed view of term recurrence across the analyzed studies.

The predominance of approaches based on machine learning and deep learning indicates potential for transferring these methodologies to other sectors such as healthcare, food security, and mining. The strong presence of terms associated with hyperspectral data and feature extraction opens opportunities for their application in new geographic regions with strategic crops, expanding the validity of the models. Moreover, the conceptual integration with environmental variables allows these techniques to be projected into different temporal scenarios, including future risk prediction. Finally, the consolidation of these trends highlights the need to strengthen capacities in data science and remote sensing across business, academic, and governmental environments to accelerate the adoption of these technologies.

RQ4: *How are studies on the application of computer vision systems for detecting heavy metals in export crops with hyperspectral imaging and XRF-IoT sensors distributed across the different quartiles of scientific journals?*

Figure 9 and Tables 11 and 12 show the distribution of articles by quartile, indexing source, and bibliometric impact. These results make it possible to assess not only the quality of the publications but also the influence of the main databases in consolidating this research field.

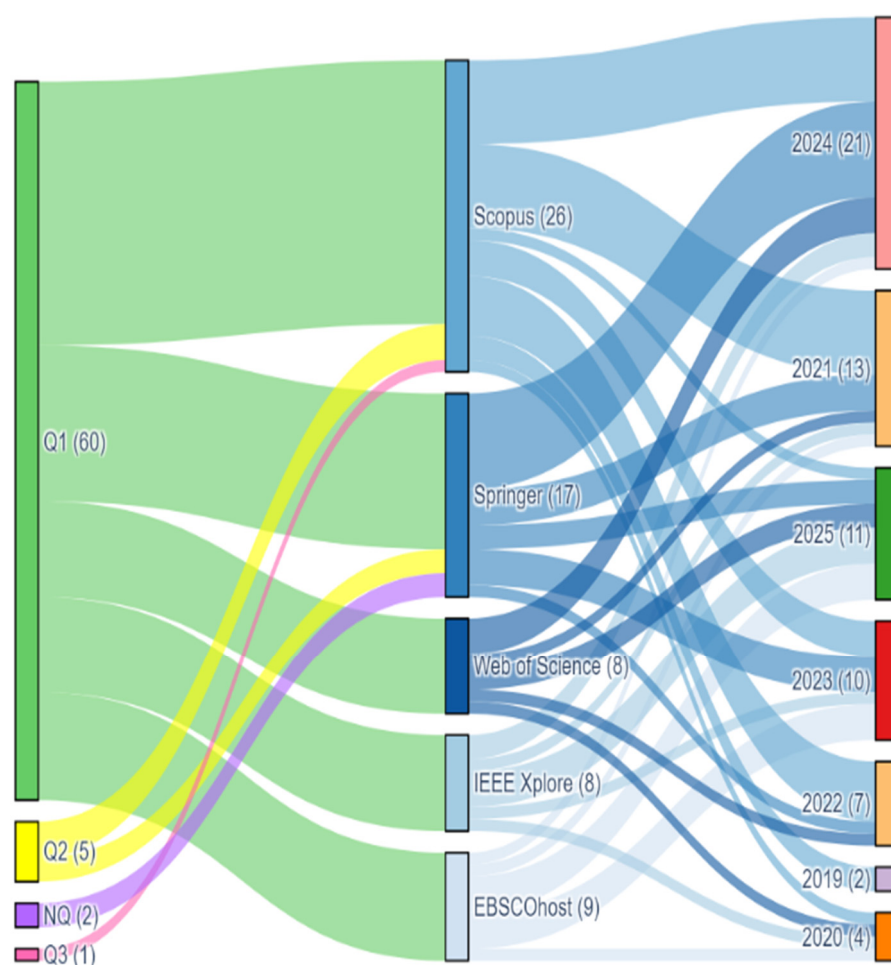


Figure 9. Publications by quartile, source, and year.

Table 11. Papers by quartile and source.

Source	NQ	Q1	Q2	Q3	Total
Scopus	0	22	3	1	26
Springer	2	13	2	0	17
EBSCOhost	0	9	0	0	9
IEEE Xplore	0	8	0	0	8
Web of Science	0	8	0	0	8
Total	2	60	5	1	68

Table 12. Impact of sources on publications.

Source	N° Papers	N° Citations	H-Index	Citations/Year	Citations/Paper
Scopus	26	531	7010	118	20
Springer Link	17	228	1004	70	13
EBSCOhost	9	439	2033	112	49
IEEE Xplore	8	167	1716	37	21
Web of Science	8	138	1444	32	17
Total	68	1503	13,207	369	22

The data reflect a clear predominance of publications in Q1 journals, with 60 papers (88.2%), evidencing the high quality and visibility of research in the area. In smaller proportions, Q2 journals account for 5 papers (7.4%) and Q3 journals for only 1 paper (1.4%), while 2 works (2.9%) were published in non-quartile journals (NQ). Regarding sources, Scopus hosts 26 papers, of which 22 correspond to Q1, confirming its role as the leading repository. Springer Link contributes 17 papers, 13 of them in Q1, while EBSCOhost, IEEE Xplore, and Web of Science contribute 9, 8, and 8 papers, respectively, all with a predominance of Q1. From an impact perspective, Scopus leads with 531 citations (20 per paper) and a cumulative H-index of 7010, followed by EBSCOhost with 439 citations (49 per paper) and a cumulative H-index of 2033. This pattern demonstrates that the most relevant research is not only concentrated in high-quartile journals but also in sources that ensure international visibility. Overall, the distribution confirms that the field develops within highly prestigious academic environments, consolidating its scientific validity.

Beyond the descriptive distribution of quartiles and databases, the evidence reveals a structural pattern in how scientific visibility is concentrated. The dominance of Q1 journals (88.2%) shows that research on CVS, HSI, and XRF-IoT is published mainly in high-impact venues but also exposes a geographic and editorial bias favoring countries with strong publication infrastructures and English-language databases. This limits the representativeness of findings in regions where export-crop contamination is most critical. Although Scopus and Springer lead in citations and H-index, the field has not expanded toward regional or sector-specific journals that could improve applicability. Overall, the distribution indicates strong scientific credibility but a narrowly focused dissemination ecosystem that should broaden to ensure greater contextual relevance.

Hasan and colleagues [98] point out that most of their reviews are concentrated in Q2 and Q1 journals, confirming that the upper quartiles are the main dissemination venues in the field of remote sensing. In line with this trend, the results of our study reveal a clear concentration in first-quartile journals (Q1: 88% of the papers), with much lower participation in Q2 and Q3. Moreover, while Hasan and colleagues provide a general overview, this research extends the analysis to a highly specialized subfield: the detection of heavy metals in export crops through computer vision. This pioneering approach not only corroborates the importance of Q1 as the primary vehicle of dissemination but also

highlights the leading role of databases such as Scopus and Springer, whose bibliometric impact reinforces the consolidation of this emerging domain.

The predominance of Q1 publications ensures the credibility and transferability of results to business sectors such as agribusiness, food security, and mining, which demand high scientific standards. The diversification of sources suggests the need to expand these studies to other databases and geographic contexts, in order to broaden dissemination in regions that remain underrepresented. The temporal evolution of these publications will allow for the detection of new research hubs and the measurement of their impact across different periods. Finally, the concentration in high-impact journals can serve as a reference to guide funding policies and international collaboration, strengthening the link between academic research and practical application across multiple sectors.

RQ5: Which countries show the highest levels of co-occurrence and cooperation in scientific publications related to computer vision systems and their application in the detection of heavy metals in export crops using hyperspectral imaging and XRF-IoT sensors?

Figure 10 and Table 13 present the scientific collaboration networks and country-level impact in studies on computer vision systems applied to the detection of heavy metals in export crops. This evidence makes it possible to identify both the main production hubs and the international cooperation networks that strengthen knowledge dissemination.

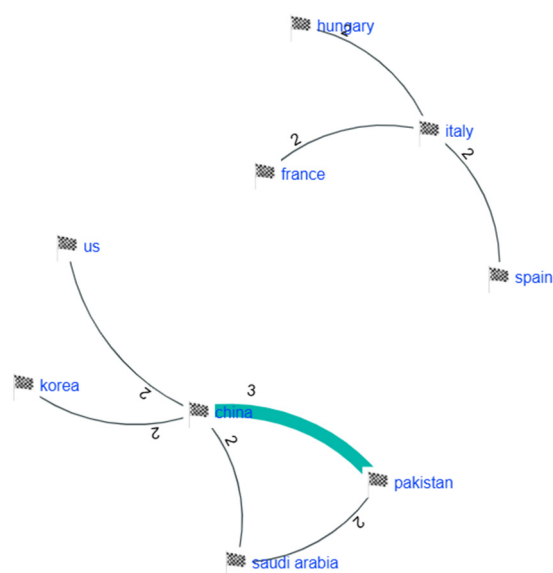


Figure 10. Bibliometric network by country.

The results show that China leads with 24 publications and 613 citations (25.5 citations per paper), also standing out as a central hub of collaboration with Pakistan (3 links), Korea, Saudi Arabia, and the United States (2 links each). The United States, with 9 publications and 122 citations (13.6 citations per paper), represents another relevant pole, although with lower connection density. India appears with 6 papers and 154 citations (25.7 citations per paper), confirming its emerging role in the field. In Europe, countries such as Italy (57 citations, 9.5 citations per paper), France (47 citations, 15.7 citations per paper), and Spain (81 citations, 40.5 citations per paper) demonstrate active collaboration, with Hungary and Belgium also standing out for their relative productivity and medium impact. Australia, with 4 papers and 198 citations (49.5 citations per paper), is noteworthy for positioning itself as a reference in quality rather than volume. Overall, the network reveals the existence of central nodes (China and Italy) and bilateral connections of varying intensity that consolidate international cooperation.

Table 13. Impact of countries on publications.

Country	Avg H-Index	Total Papers	No. of Citations	Citations/Paper	PY Start
China	180.8	24	613	25.5	2019
US	204.8	9	122	13.6	2018
India	142.3	6	154	25.7	2019
Italy	320.5	6	57	9.5	2019
Saudi Arabia	189.0	6	28	4.7	2020
Australia	179.3	4	198	49.5	2018
Canada	167.8	4	0	0.0	2019
Egypt	136.0	3	14	4.7	2022
France	331.7	3	47	15.7	2020
Hungary	301.0	3	20	6.7	2021
Korea	154.3	3	24	8.0	2019
Pakistan	150.3	3	19	6.3	2019
Spain	220.3	3	26	8.7	2019
Belgium	307.0	2	81	40.5	2021
Germany	204.5	2	27	13.5	2019
Greece	99.0	2	1	0.5	2019
Morocco	217.0	2	36	18.0	2019
...	...	...	...	...	...
Total	196.1	101	1838	18.2	

In line with our findings, R. Bertoglio and colleagues [22] argue that the United States and China concentrate the highest volumes of scientific production, with more than 600 publications each, reinforcing their role as global leaders. Complementarily, Andrade-Mogollón and his research team [23] identify frequent collaborations between Saudi Arabia, Pakistan, and the United Kingdom, which confirms our results. Likewise, Sabour and collaborators [84] highlight China as the most cooperative country (37 links), followed by the United States (27) and Germany (19), consistent with our bibliometric network, where China emerges as a central hub of cooperation with Pakistan, Korea, Saudi Arabia, and the United States. Finally, Hasan and his research group [98] recognize China as one of the most influential co-authors, which is corroborated in this study by its leadership with 24 publications and 613 citations, consolidating its position as a central collaboration node.

The leadership of China and the U.S. indicates the potential to transfer these experiences to strategic business sectors such as agribusiness, mining, and environmental health, leveraging their capacity to generate applied knowledge. The strong participation of European and Oceanian countries suggests the importance of extending these collaborations to underrepresented regions such as Latin America or sub-Saharan Africa. Temporal analysis will make it possible to evaluate how these networks evolve in future scenarios, particularly in light of new international regulatory requirements. Finally, the cooperation dynamics identified here can serve as a reference for promoting interregional consortia that integrate academic research and practical applications in diverse productive contexts.

RQ6: *What central thematic categories are identified in research integrating computer vision systems with the detection of heavy metals in export crops using hyperspectral imaging and XRF-IoT sensors?*

Figure 11 and Table 14 show the emerging topics obtained from the keyword analysis, considering two parameters: centrality (degree of relevance, horizontal axis) and density (degree of development, vertical axis). These results make it possible to identify the positioning of categories within the conceptual structure of the field.

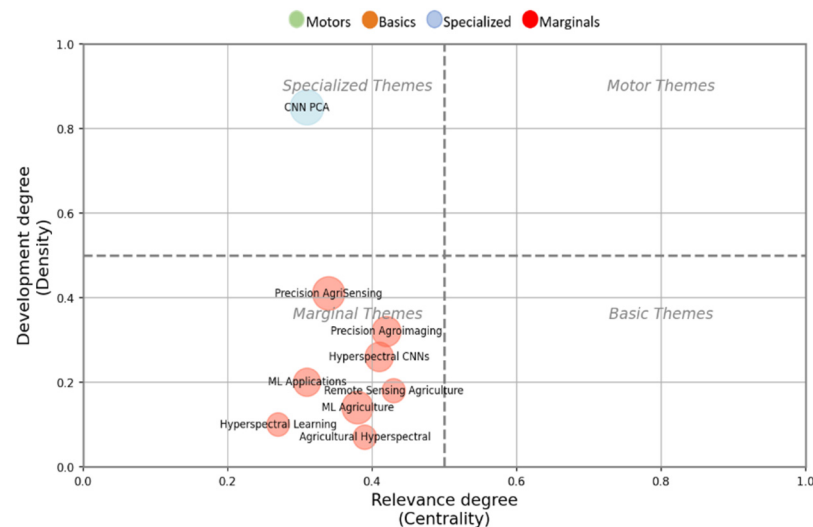


Figure 11. Thematic map on research areas.

Table 14. Topic density, centrality, total citations, and categories.

Topic	Density	Centrality	Total Citations	Category
CNN PCA	0.85	0.31	905	Specialized
Precision AgriSensing	0.41	0.34	825	Marginal
Precision Agroimaging	0.32	0.42	533	Marginal
Hyperspectral CNNs	0.26	0.41	638	Marginal
ML Applications	0.20	0.31	467	Marginal
Remote Sensing Agriculture	0.18	0.43	1142	Marginal
ML Agriculture	0.14	0.38	779	Marginal
Hyperspectral Learning	0.10	0.27	1076	Marginal
Agricultural Hyperspectral	0.07	0.39	1160	Marginal

The thematic analysis was conducted from a structural science-mapping perspective based on centrality–density relationships derived from the Callon strategic diagram, enabling an interpretive understanding of the intellectual organization of the field rather than a purely descriptive bibliometric classification.

The density (D) of a theme measures its internal cohesion, determined by the intensity of connections among the keywords that constitute it, whereas centrality (C) captures the degree of interaction between that theme and other clusters within the network. These indicators are calculated according to the following expressions:

$$D_i = 100 \times \frac{\sum e_{jk}}{n_i(n_i-1)/2}$$

$$C_i = 10 \times \sum e_{ij}$$

where e_{jk} denotes the connection strength between keywords j and k within cluster i , n_i represents the total number of keywords in cluster i , and e_{ij} indicates the connection strength between cluster i and any other cluster j .

The findings reveal that CNN PCA is positioned as a specialized topic, with high density (0.85) and a total of 905 citations, reflecting a methodologically robust field but with limited connections to other topics. Most themes are concentrated in the marginal category, characterized by moderate relevance and low development, although with a significant number of citations, which indicates growing interest. Within this group, Remote Sensing Agriculture (centrality 0.43, 1142 citations) and Agricultural Hyperspectral (centrality 0.39, 1160 citations) stand out, as despite being classified as marginal, they

accumulate high influence in the literature. Precision AgriSensing (825 citations), Precision Agroimaging (533 citations), and Hyperspectral CNNs (638 citations) also emerge, pointing to the integration of computer vision approaches with hyperspectral techniques applied to precision agriculture. Other topics, such as ML Applications, ML Agriculture, and Hyperspectral Learning, although showing lower density, represent lines of research in consolidation that articulate machine learning and remote sensing. Overall, the distribution evidences a field with a solid methodological core and a set of emerging agricultural topics in the process of conceptual maturation.

In line with our findings, R. Bertoglio and collaborators [22] identify remote sensing and the Internet of Things (IoT) as driving themes, while Artificial Intelligence (AI) emerges as a development axis with strong projection. Similarly, Botero-Valencia and colleagues [24] point out that AI and remote sensing are consolidated as central areas, while IoT and neural networks are considered emerging.

These results suggest that future evolution should focus on strengthening marginal topics with greater citation impact, transforming them into central axes of research. The consolidation of topics such as remote sensing agriculture or agricultural hyperspectral could have direct applications in sectors such as food security, mining, and environmental management. Likewise, comparative analysis across different geographic regions and periods will make it possible to validate the relevance of these categories in diverse productive contexts. Finally, the findings encourage fostering interdisciplinary projects that integrate AI, precision agriculture, and advanced sensing technologies, accelerating the transfer of these innovations to business and governmental sectors.

5. Conclusions and Future Research

Regarding RQ1, the findings indicate that the evaluation of computer vision systems in the analyzed studies predominantly relies on statistical error metrics such as RMSE and MAE, reflecting a consistent methodological preference within the reviewed literature. The limited use of complementary indicators—such as inference time or ROC curves—suggests that performance assessment across the 68 papers primarily emphasizes numerical accuracy rather than robustness or operational behavior. These observations are confined to the analyzed corpus and should not be generalized beyond the scope of the reviewed studies.

With respect to RQ2, Python emerges as the predominant programming language in the selected papers, mainly due to its strong ecosystem for AI and computer vision. MATLAB and Scala maintain relevance in academic or simulation-oriented contexts, while C++, Java, and Swift appear only marginally. These trends describe the technological choices of the 68 studies and do not represent the full range of tools used in the broader research community.

In relation to RQ4, most of the reviewed papers were published in Q1 journals, indicating that the studies captured by our search strategy are concentrated in high-impact venues. Publications in Q2 and Q3 are less frequent, while Q4 and non-indexed sources appear only exceptionally. This pattern reflects the distribution of the 68 included studies and cannot be extrapolated to the entire global publication landscape on the topic.

Regarding RQ5, the bibliometric patterns observed show that China and the United States contribute the largest number of papers within the analyzed corpus, followed by India, various European countries, and Australia. These findings describe only the geographic distribution present in the 68 studies and do not claim to represent worldwide research activity. Nonetheless, the presence of multi-country collaborations in this sample suggests that cooperation networks play a relevant role in advancing the field.

Finally, with respect to RQ6, the thematic analysis indicates that CNN PCA appears as a specialized cluster in the included studies, while areas such as remote sensing agriculture

and agricultural hyperspectral show high citation activity. These patterns reflect the conceptual organization found exclusively within the 68 reviewed papers. Future research could strengthen marginal themes and further integrate AI methods with hyperspectral and XRF-IoT technologies to improve methodological robustness, operational reliability, and real-world applicability.

Looking ahead, it is recommended to deepen the integration of operational and statistical metrics, diversify programming languages beyond Python to guarantee technological scalability, and promote publication strategies in more accessible repositories that complement the predominance of Q1 journals. Likewise, it is a priority to strengthen international cooperation with underrepresented regions to broaden global impact and consolidate marginal topics with higher citation impact, so that they evolve into central and sustainable research cores.

Future research should prioritize the improvement of calibration and standardization protocols for hyperspectral imaging and XRF-IoT systems, developing models that remain robust under environmental variability, and validating multi-sensor platforms in real field conditions rather than controlled laboratory settings. It is also necessary to advance the creation of open-access datasets that enhance reproducibility and allow fair benchmarking of emerging models across crops and regions. Integrating HSI and XRF-IoT measurements with edge computing and AI-driven decision-support tools represents a promising direction to strengthen food safety and increase the competitiveness of export crops at a global scale.

From an overall maturity perspective, current CVS-, HSI-, and XRF-IoT-based approaches for heavy-metal detection in export crops can be considered to be at an intermediate stage of technological readiness. While substantial progress has been achieved in spectral-spatial modeling accuracy and laboratory validation, several critical gaps persist that limit large-scale operational deployment. These include the lack of standardized and openly available multi-site datasets, insufficient normalization of evaluation protocols, limited reporting of computational cost and real-time feasibility, and the absence of widely adopted communication and interoperability standards for XRF-IoT agricultural sensing architectures. Moreover, architectural integration between HSI analytics and in situ elemental sensing remains mostly experimental and rarely validated under field conditions. Consequently, although the technological foundations are solid, the transition toward fully scalable, interoperable, and industry-ready heavy-metal monitoring systems in export-crop environments is still emerging and requires coordinated advances in dataset standardization, edge-enabled architectures, and robust IoT communication frameworks.

Author Contributions: Conceptualization, J.G.-C., K.T.-A., A.S.-M. and K.O.-M.; methodology, J.G.-C. and R.M.M.; software, K.T.-A., C.D.-V.-J. and A.N.M.; validation, K.T.-A., K.O.-M., A.S.-M. and C.D.-V.-J.; investigation, J.G.-C., C.D.-V.-J. and R.M.M.; resources, R.M.M.; writing—original draft preparation, K.T.-A., A.S.-M. and K.O.-M.; writing—review and editing, J.G.-C., C.D.-V.-J. and A.N.M.; project administration, J.G.-C., A.S.-M. and A.N.M. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by the Universidad Nacional Mayor de San Marcos.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: This study is a systematic review; therefore, no new datasets were generated or analyzed. All data supporting the findings are derived from publicly available sources cited in the manuscript.

Conflicts of Interest: The authors declare no conflicts of interest. The funders had no role in the design of the study; in the collection, analyses, or interpretation of data; in the writing of the manuscript; or in the decision to publish the results.

Appendix A

Table A1. Search Equations by Source.

Source	Search Equation
Web of Science	("artificial vision" OR "computer vision" OR "image analysis" OR "machine vision" OR "deep learning" OR "neural networks" OR "convolutional networks" OR "artificial intelligence" OR "machine learning" OR "transfer learning" OR "image classification" OR "object detection" OR "semantic segmentation") AND ((("heavy metals" OR pollutants OR "toxic metals" OR "metal traces" OR hyperspectral OR "metal ions" OR "metal toxicity" OR "mineral residues" OR "heavy elements" OR xrf) AND ("agriculture 4.0" OR "agricultural iot" OR "precision agriculture" OR agriculture)) (Title) or ("artificial vision" OR "computer vision" OR "image analysis" OR "machine vision" OR "deep learning" OR "neural networks" OR "convolutional networks" OR "artificial intelligence" OR "machine learning" OR "transfer learning" OR "image classification" OR "object detection" OR "semantic segmentation") AND ((("heavy metals" OR pollutants OR "toxic metals" OR "metal traces" OR hyperspectral OR "metal ions" OR "metal toxicity" OR "mineral residues" OR "heavy elements" OR xrf) AND ("agriculture 4.0" OR "agricultural iot" OR "precision agriculture" OR agriculture)) (Author Keywords)
Scopus	(KEY(("artificial vision" OR "computer vision" OR "image analysis" OR "machine vision" OR "deep learning" OR "neural networks" OR "convolutional networks" OR "artificial intelligence" OR "machine learning" OR "transfer learning" OR "image classification" OR "object detection" OR "semantic segmentation") AND ((("heavy metals" OR pollutants OR "toxic metals" OR "metal traces" OR hyperspectral OR "metal ions" OR "metal toxicity" OR "mineral residues" OR "heavy elements" OR xrf) AND ("agriculture 4.0" OR "agricultural iot" OR "precision agriculture")) OR TITLE(("artificial vision" OR "computer vision" OR "image analysis" OR "machine vision" OR "deep learning" OR "neural networks" OR "convolutional networks" OR "artificial intelligence" OR "machine learning" OR "transfer learning" OR "image classification" OR "object detection" OR "semantic segmentation") AND ((("heavy metals" OR pollutants OR "toxic metals" OR "metal traces" OR hyperspectral OR "metal ions" OR "metal toxicity" OR "mineral residues" OR "heavy elements" OR xrf) AND ("agriculture 4.0" OR "agricultural iot" OR "precision agriculture")))))
IEEE Xplore	("Author Keywords": "artificial vision" OR "Author Keywords": "computer vision" OR "Author Keywords": "image analysis" OR "Author Keywords": "machine vision" OR "Author Keywords": "deep learning" OR "Author Keywords": "neural networks" OR "Author Keywords": "convolutional networks" OR "Author Keywords": "artificial intelligence" OR "Author Keywords": "machine learning" OR "Author Keywords": "transfer learning" OR "Author Keywords": "image classification" OR "Author Keywords": "object detection" OR "Author Keywords": "semantic segmentation") AND ("Author Keywords": "heavy metals" OR "Author Keywords": "pollutants" OR "Author Keywords": "toxic metals" OR "Author Keywords": "metal traces" OR "Author Keywords": "hyperspectral" OR "Author Keywords": "metal ions" OR "Author Keywords": "metal toxicity" OR "Author Keywords": "mineral residues" OR "Author Keywords": "heavy elements" OR "Author Keywords": "xrf" AND ("Author Keywords": "agriculture 4.0" OR "Author Keywords": "agricultural iot" OR "Author Keywords": "precision agriculture") OR ("Document Title": "artificial vision" OR "Document Title": "computer vision" OR "Document Title": "image analysis" OR "Document Title": "machine vision" OR "Document Title": "deep learning" OR "Document Title": "neural networks" OR "Document Title": "convolutional networks" OR "Document Title": "artificial intelligence" OR "Document Title": "machine learning" OR "Document Title": "transfer learning" OR "Document Title": "image classification" OR "Document Title": "object detection" OR "Document Title": "semantic segmentation") AND ("Document Title": "heavy metals" OR "Document Title": "pollutants" OR "Document Title": "toxic metals" OR "Document Title": "metal traces" OR "Document Title": "hyperspectral" OR "Document Title": "metal ions" OR "Document Title": "metal toxicity" OR "Document Title": "mineral residues" OR "Document Title": "heavy elements" OR "Document Title": "xrf" AND ("Document Title": "agriculture 4.0" OR "Document Title": "agricultural iot" OR "Document Title": "precision agriculture") OR ("Abstract": "artificial vision" OR "Abstract": "computer vision" OR "Abstract": "image analysis" OR "Abstract": "machine vision" OR "Abstract": "deep learning" OR "Abstract": "neural networks" OR "Abstract": "convolutional networks" OR "Abstract": "artificial intelligence" OR "Abstract": "machine learning" OR "Abstract": "transfer learning" OR "Abstract": "image classification" OR "Abstract": "object detection" OR "Abstract": "semantic segmentation") AND ("Abstract": "heavy metals" OR "Abstract": "pollutants" OR "Abstract": "toxic metals" OR "Abstract": "metal traces" OR "Abstract": "hyperspectral" OR "Abstract": "metal ions" OR "Abstract": "metal toxicity" OR "Abstract": "mineral residues" OR "Abstract": "heavy elements" OR "Abstract": "xrf" AND ("Abstract": "agriculture 4.0" OR "Abstract": "agricultural iot" OR "Abstract": "precision agriculture")

Table A1. Cont.

Source	Search Equation
EBSCOhost	TI (("artificial vision" OR "computer vision" OR "image analysis" OR "machine vision" OR "deep learning" OR "neural networks" OR "convolutional networks" OR "artificial intelligence" OR "machine learning" OR "transfer learning" OR "image classification" OR "object detection" OR "semantic segmentation")) AND TI (("heavy metals" OR pollutants OR "toxic metals" OR "metal traces" OR hyperspectral OR "metal ions" OR "metal toxicity" OR "mineral residues" OR "heavy elements" OR xrf)) AND TI (("agriculture 4.0" OR "agricultural iot" OR "precision agriculture" OR agriculture)) OR SU (("artificial vision" OR "computer vision" OR "image analysis" OR "machine vision" OR "deep learning" OR "neural networks" OR "convolutional networks" OR "artificial intelligence" OR "machine learning" OR "transfer learning" OR "image classification" OR "object detection" OR "semantic segmentation")) AND SU (("heavy metals" OR pollutants OR "toxic metals" OR "metal traces" OR hyperspectral OR "metal ions" OR "metal toxicity" OR "mineral residues" OR "heavy elements" OR xrf)) AND SU (("agriculture 4.0" OR "agricultural iot" OR "precision agriculture" OR agriculture)) OR AB (("artificial vision" OR "computer vision" OR "image analysis" OR "machine vision" OR "deep learning" OR "neural networks" OR "convolutional networks" OR "artificial intelligence" OR "machine learning" OR "transfer learning" OR "image classification" OR "object detection" OR "semantic segmentation")) AND AB (("heavy metals" OR "metal ions" OR "metal toxicity" OR "mineral residues" OR "heavy elements" OR xrf)) AND AB (("agriculture 4.0" OR "agricultural iot" OR "precision agriculture" OR agriculture))
Springer Nature Link	("artificial vision" OR "computer vision" OR "image analysis" OR "machine vision" OR "deep learning" OR "neural networks" OR "convolutional networks" OR "artificial intelligence" OR "machine learning" OR "transfer learning" OR "image classification" OR "object detection" OR "semantic segmentation") AND ("heavy metals" OR pollutants OR "toxic metals" OR "metal traces" OR hyperspectral OR "metal ions" OR "metal toxicity" OR "mineral residues" OR "heavy elements" OR xrf) AND ("agriculture 4.0" OR "agricultural iot" OR "precision agriculture" OR agriculture))

Appendix B

Table A2. Consolidated Matrix of Results After Applying the Exclusion Criteria.

Information Source	Initial Results (N)	Results After 1st Filter (EC 1–2)	Results After 2nd Filter (EC 3–4)	Results After 3rd Filter (EC 5–6)	Results After 4th Filter (EC 7–8)	% Final
Web of Science	74	66	42	14	9	12
Scopus	649	425	131	30	26	37
IEEE Xplore	149	134	16	10	9	12
EBSCOhost	743	477	318	12	9	12
Spring Nature Link	10,105	7805	1309	21	19	27
Total	11,720	8907	1816	87	72	100

Appendix C

Table A3. Sensor-level characterization of studies directly estimating heavy metals using HSI/Vis–NIR spectroscopy and XRF-related evidence in the primary set.

Ref.	Environment	Platform	Sensor/Instrument	Spectral Range	Spatial Resolution	Target Metals /Analytes	Model + Performance (as Reported)
[34]	Soil (arable fields)	Laboratory	Vis-NIR + XRF spectra fusion (no IoT)	Vis-NIR 350–2500 nm; XRF energía 1–115 keV	NR	Soil attributes (p.ej., Ca y otros; no metales como objetivo principal)	PLS + CNN con fusión (SF-PLS / SF-CNN); desempeño reportado para atributos
[46]	Soil (farmland)	Satellite	Hyperspectral sensor AHSI on ZY-1-02D	400–2500 nm (166 channels; ~10 nm)	30 m	Mn, Cu, Zn, Pb, Cr y Ni (soil)	Stacked AdaBoost; reports R ² (e.g., Cr ≈ 0.48, Cu ≈ 0.49, As ≈ 0.48 en el texto extraído)
[52]	Soil	Laboratory	Vis-NIR reflectance (SR + FR spectrometers; FR fibered ARCOptix)	SR 400–1100 nm (2.5 nm) + FR 400–2500 nm (8 nm)	NR	Ni, Co, Cd, Pb, Cu	ML on Vis-NIR; reporta desempeño con R ² /RMSE (detallado en el paper)

Table A3. Cont.

Ref.	Environment	Platform	Sensor/Instrument	Spectral Range	Spatial Resolution	Target Metals /Analytes	Model + Performance (as Reported)
[75]	Agricultural soil (Eastern Junggar coalfield)	Laboratory spectroscopy	ASD FieldSpec3 Spectrometer (Analytical Spectral Devices)	400–2500 nm (Vis-NIR-SWIR)	NR (point measurement)	Cr—soil contamination classification (polluted vs. non-polluted)	Data augmentation + DNN: Accuracy 96.25%; DA-SVM (poly kernel): 95.61%; DA-KNN (k = 16): 95.62%; Sensitivity 82.35%, Specificity 97.20%
[67]	Soil (topsoil)	Laboratory spectroscopy	FOSS XDS Rapid Content Analyzer (VNIR spectra) + chemical assay	400–2500 nm (el PDF muestra “resolution = 0.5 m”; unidad no se renderiza bien)	Spectral sampling ~0.5 nm	As	Deep learning/ML; reporta R ² , MSE (y métricas como AUC aparecen en el texto)
[56]	Soil (agricultural)	Laboratory	Vis-NIR reflectance spectroscopy	400–2400 nm	~3 nm (350–1000), ~10 nm (1000–2500)	Hg, Cr, Cu	Fractional-order derivative + modelos (p.ej., GRNN/PLSR/RF según el paper); reporta R ² /RMSE

References

1. Agrawal, N.; Govil, H. A deep residual convolutional neural network for mineral classification. *Adv. Space Res.* **2023**, *71*, 3186–3202. [\[CrossRef\]](#)
2. Ashraf, M.; Chen, L.; Innab, N.; Umer, M.; Baili, J.; Kim, T.-H.; Ashraf, I. Novel 3-D deep neural network architecture for crop classification using remote sensing-based hyperspectral images. *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.* **2024**, *17*, 12649–12665. [\[CrossRef\]](#)
3. Ayuba, D.L.; Guillemaut, J.-Y.; Marti-Cardona, B.; Mendez, O. HyperKon: A self-supervised contrastive network for hyperspectral image analysis. *Remote Sens.* **2024**, *16*, 3399. [\[CrossRef\]](#)
4. Hassan, S.I.; Alam, M.M.; Illahi, U.; Al Ghamdi, M.A.; Almotiri, S.H.; Su’ud, M.M. A systematic review on monitoring and advanced control strategies in smart agriculture. *IEEE Access* **2021**, *9*, 32517–32548. [\[CrossRef\]](#)
5. Yu, K.; Zhong, M.; Zhu, W.; Rashid, A.; Han, R.; Virk, M.S.; Duan, K.; Zhao, Y.; Ren, X. Advances in computer vision and spectroscopy techniques for non-destructive quality assessment of citrus fruits: A comprehensive review. *Foods* **2025**, *14*, 386. [\[CrossRef\]](#)
6. Guerri, M.F.; Distant, C.; Spagnolo, P.; Bougourzi, F.; Taleb-Ahmed, A. Deep learning techniques for hyperspectral image analysis in agriculture: A review. *ISPRS Open J. Photogramm. Remote Sens.* **2024**, *12*, 100062. [\[CrossRef\]](#)
7. Alajmi, M.; Mengash, H.A.; Eltahir, M.M.; Assiri, M.; Ibrahim, S.S.; Salama, A.S. Exploiting hyperspectral imaging and optimal deep learning for crop type detection and classification. *IEEE Access* **2023**, *11*, 124985–124995. [\[CrossRef\]](#)
8. Bourriz, M.; Hajji, H.; Laamrani, A.; Elbounani, N.; Abdelali, H.A.; Bourzeix, F.; El-Battay, A.; Amazirh, A.; Chehbouni, A. Integration of hyperspectral imaging and AI techniques for crop type mapping: Present status, trends, and challenges. *Remote Sens.* **2025**, *17*, 1574. [\[CrossRef\]](#)
9. Yele, V.P.; Sedamkar, R.R.; Alegavi, S. Systematic analysis of effective segmentation and classification for land use land cover in hyperspectral image using deep learning methods: A review of the state of the art. In *2024 20th CSI International Symposium on Artificial Intelligence and Signal Processing (AISP)*; IEEE: New York, NY, USA, 2024; pp. 1–8.
10. Amirgaliyev, Y.; Mukhamediev, R.; Merembayev, T.; Kuchin, Y.; Ataniyazova, A.; Omarova, P. Remote sensing and machine learning algorithms to predict soil salinity in southern Kazakhstan. *Discov. Sustain.* **2024**, *5*, 363. [\[CrossRef\]](#)
11. Brom, J.; Duffková, R.; Haberle, J.; Zajíček, A.; Nedbal, V.; Bernasová, T.; Křováková, K. Identification of infiltration features and hydraulic properties of soils based on crop water stress derived from remotely sensed data. *Remote Sens.* **2021**, *13*, 4127. [\[CrossRef\]](#)
12. Chen, S.; Lou, F.; Tuo, Y.; Tan, S.; Peng, K.; Zhang, S.; Wang, Q. Prediction of soil water content based on hyperspectral reflectance combined with competitive adaptive reweighted sampling and random frog feature extraction and the back-propagation artificial neural network method. *Water* **2023**, *15*, 2726. [\[CrossRef\]](#)
13. Bang, G.-H.; Gwon, N.-H.; Cho, M.-J.; Park, J.-Y.; Baek, S.-S. Developing a real-time water quality simulation toolbox using machine learning and application programming interface. *J. Environ. Manag.* **2025**, *377*, 124719. [\[CrossRef\]](#) [\[PubMed\]](#)
14. Chen, J.; Zhao, L.; Wang, B.; He, X.; Duan, L.; Yu, G. Uncovering global risk to human and ecosystem health from pesticides in agricultural surface water using a machine learning approach. *Environ. Int.* **2024**, *194*, 109154. [\[CrossRef\]](#)

15. Fei, X.; Lou, Z.; Sheng, M.; Xiaonan, L.; Ren, Z.; Xiao, R. Source-oriented stochastic health risk assessment of toxic metals in soil via a hybrid model and Monte Carlo simulation. *Ecotoxicol. Environ. Saf.* **2024**, *286*, 117209. [[CrossRef](#)] [[PubMed](#)]
16. Casado-García, A.; Heras, J.; Milella, A.; Marani, R. Semi-supervised deep learning and low-cost cameras for the semantic segmentation of natural images in viticulture. *Precis. Agric.* **2022**, *23*, 2001–2026. [[CrossRef](#)]
17. Fahey, T.; Pham, H.; Gardi, A.; Sabatini, R.; Stefanelli, D.; Goodwin, I.; Lamb, D.W. Active and passive electro-optical sensors for health assessment in food crops. *Sensors* **2020**, *21*, 171. [[CrossRef](#)]
18. Swaraj, S.; Aparna, S. An empirical overview for plant disease detection and severity level classification using hyperspectral images. In *2024 Eighth International Conference on Parallel, Distributed and Grid Computing (PDGC)*; IEEE: New York, NY, USA, 2025; pp. 182–189.
19. Dash, S.; Chakravarty, S.; Giri, N.C.; Agyekum, E.B.; AboRas, K.M. Minimum noise fraction and long short-term memory model for hyperspectral imaging. *Int. J. Comput. Intell. Syst.* **2024**, *17*, 16. [[CrossRef](#)]
20. Dhaliwal, J.K.; Panday, D.; Robertson, G.P.; Saha, D. Machine learning reveals dynamic controls of soil nitrous oxide emissions from diverse long-term cropping systems. *J. Environ. Qual.* **2025**, *54*, 132–146. [[CrossRef](#)]
21. Droz, B.; Payraudeau, S.; Rodríguez Martín, J.A.; Tóth, G.; Panagos, P.; Montanarella, L.; Borrelli, P.; Imfeld, G. Copper content and export in European vineyard soils influenced by climate and soil properties. *Environ. Sci. Technol.* **2021**, *55*, 7327–7334. [[CrossRef](#)]
22. Bertoglio, R.; Corbo, C.; Renga, F.M.; Matteucci, M. The digital agricultural revolution: A bibliometric analysis literature review. *IEEE Access* **2021**, *9*, 134762–134782. [[CrossRef](#)]
23. Andrade-Mogollon, T.; Gamboa-Cruzado, J.; Amayo-Gamboa, F. Systematic literature review of generative AI and IoT as key technologies for precision agriculture. *Comput. Sist.* **2025**, *29*, 857–882. [[CrossRef](#)]
24. Botero-Valencia, J.; García-Pineda, V.; Valencia-Arias, A.; Valencia, J.; Reyes-Vera, E.; Mejia-Herrera, M.; Hernández-García, R. Machine learning in sustainable agriculture: Systematic review and research perspectives. *Agriculture* **2025**, *15*, 377. [[CrossRef](#)]
25. Ang, K.L.-M.; Seng, J.K.P. Big data and machine learning with hyperspectral information in agriculture. *IEEE Access* **2021**, *9*, 36699–36718. [[CrossRef](#)]
26. Barbedo, J.G.A. A review on the combination of deep learning techniques with proximal hyperspectral images in agriculture. *Comput. Electron. Agric.* **2023**, *210*, 107920. [[CrossRef](#)]
27. Feinberg, A.; Stenke, A.; Peter, T.; Hinckley, E.-L.S.; Driscoll, C.T.; Winkel, L.H.E. Reductions in the deposition of sulfur and selenium to agricultural soils pose risk of future nutrient deficiencies. *Commun. Earth Environ.* **2021**, *2*, 101. [[CrossRef](#)]
28. Haagsma, M.; Hagerty, C.H.; Kroese, D.R.; Selker, J.S. Detection of soil-borne wheat mosaic virus using hyperspectral imaging: From lab to field scans and from hyperspectral to multispectral data. *Precis. Agric.* **2023**, *24*, 1030–1048. [[CrossRef](#)]
29. Wang, S. Effectiveness of traditional augmentation methods for rebar counting using UAV imagery with Faster R-CNN and YOLOv10-based transformer architectures. *Sci. Rep.* **2025**, *15*, 33702. [[CrossRef](#)] [[PubMed](#)]
30. Hsieh, T.-H.; Kiang, J.-F. Comparison of CNN algorithms on hyperspectral image classification in agricultural lands. *Sensors* **2020**, *20*, 1734. [[CrossRef](#)]
31. Huan, J.; Yuan, J.; Zhang, H.; Xu, X.; Shi, B.; Zheng, Y.; Li, X.; Zhang, C.; Hu, Q.; Fan, Y.; et al. Identification of agricultural surface source pollution in plain river network areas based on 3D-EEMs and convolutional neural networks. *Water Sci. Technol.* **2024**, *89*, 1961–1980. [[CrossRef](#)]
32. Hütt, C.; Bolten, A.; Hüging, H.; Bareth, G. UAV LiDAR metrics for monitoring crop height, biomass and nitrogen uptake: A case study on a winter wheat field trial. *PFG–J. Photogramm. Remote Sens. Geoinf. Sci.* **2023**, *91*, 65–76. [[CrossRef](#)]
33. Islam, M.; Bijjahalli, S.; Fahey, T.; Gardi, A.; Sabatini, R.; Lamb, D.W. Destructive and non-destructive measurement approaches and the application of AI models in precision agriculture: A review. *Precis. Agric.* **2024**, *25*, 1127–1180. [[CrossRef](#)]
34. Javadi, S.H.; Munna, M.A.; Mouazen, A.M. Fusion of Vis-NIR and XRF spectra for estimation of key soil attributes. *Geoderma* **2021**, *385*, 114851. [[CrossRef](#)]
35. Jindo, K.; Kozan, O.; Iseki, K.; Maestrini, B.; van Evert, F.K.; Wubengeda, Y.; Arai, E.; Shimabukuro, Y.E.; Sawada, Y.; Kempenaar, C. Potential utilization of satellite remote sensing for field-based agricultural studies. *Chem. Biol. Technol. Agric.* **2021**, *8*, 58. [[CrossRef](#)]
36. Gamboa-Cruzado, J.; Menendez-Morales, C.; Del Carpio, C.F.; López-Goycochea, J.; Arévalo, A.A.; Vargas, C.R. Use of chatbots in e-commerce: A comprehensive systematic review. *J. Theor. Appl. Inf. Technol.* **2023**, *101*, 1172–1183.
37. Villavicencio, H.E.; Gamboa-Cruzado, J.; López-Goycochea, J.; Soto, L.S. The role of artificial intelligence in the diagnosis of neoplastic diseases: A systematic and bibliometric review. *Int. J. Online Biomed. Eng.* **2024**, *20*, 43–68. [[CrossRef](#)]
38. Wang, S.; Eum, I.; Park, S.; Kim, J. A labelled dataset for rebar counting inspection on construction sites using unmanned aerial vehicles. *Data Brief.* **2024**, *55*, 110720. [[CrossRef](#)]
39. Khan, S.; Majdalawieh, M.; Boufama, B.; Sharma, Y.; Basani, A. Unlocking the potential of simulated hyperspectral imaging in agro environmental analysis: A comprehensive study of algorithmic approaches. *Front. Agron.* **2024**, *6*, 1435234. [[CrossRef](#)]

40. Kishor Kumar Reddy, C.; Daduvy, A.; Madana Mohana, R.; Assiri, B.; Shuaib, M.; Alam, S.; Abdullah Sheneamer, M. Enhancing precision agriculture and land cover classification: A self-attention 3D convolutional neural network approach for hyperspectral image analysis. *IEEE Access* **2024**, *12*, 125592–125608. [\[CrossRef\]](#)
41. Kourounioti, O.; Temenos, A.; Temenos, N.; Oikonomou, E.; Doulamis, A.; Doulamis, N. UAVINE-XAI: eXplainable AI-based spectral band selection for vineyard monitoring using UAV hyperspectral data. *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.* **2025**, *18*, 10095–10104. [\[CrossRef\]](#)
42. Laroche-Pinel, E.; Vasquez, K.R.; Brillante, L. Assessing grapevine water status in a variably irrigated vineyard with NIR/SWIR hyperspectral imaging from UAV. *Precis. Agric.* **2024**, *25*, 2356–2374. [\[CrossRef\]](#)
43. Lavanya, K.; Jaya Subalakshmi, R.; Tamizharasi, T.; Jane, L.; Victor, A. Unsupervised unmixing and segmentation of hyper spectral images accounting for soil fertility. *Scalable Comput. Pract. Exp.* **2022**, *23*, 291–301. [\[CrossRef\]](#)
44. Li, K.-Y.; Sampaio de Lima, R.; Burnside, N.G.; Vahtmäe, E.; Kutser, T.; Sepp, K.; Cabral Pinheiro, V.H.; Yang, M.-D.; Vain, A.; Sepp, K. Toward automated machine learning-based hyperspectral image analysis in crop yield and biomass estimation. *Remote Sens.* **2022**, *14*, 1114. [\[CrossRef\]](#)
45. Li, J.; Ge, Y.; Puntel, L.A.; Heeren, D.M.; Bai, G.; Balboa, G.R.; Gamon, J.A.; Arkebauer, T.J.; Shi, Y. Devising optimized maize nitrogen stress indices in complex field conditions from UAV hyperspectral imagery. *Precis. Agric.* **2025**, *26*, 3. [\[CrossRef\]](#)
46. Lin, N.; Jiang, R.; Li, G.; Yang, Q.; Li, D.; Yang, X. Estimating the heavy metal contents in farmland soil from hyperspectral images based on stacked AdaBoost ensemble learning. *Ecol. Indic.* **2022**, *143*, 109330. [\[CrossRef\]](#)
47. Meng, R.; Lv, Z.; Yan, J.; Chen, G.; Zhao, F.; Zeng, L.; Xu, B. Development of spectral disease indices for southern corn rust detection and severity classification. *Remote Sens.* **2020**, *12*, 3233. [\[CrossRef\]](#)
48. Misbah, K.; Laamrani, A.; Khechba, K.; Dhiba, D.; Chehbouni, A. Multi-sensors remote sensing applications for assessing, monitoring, and mapping NPK content in soil and crops in African agricultural land. *Remote Sens.* **2021**, *14*, 81. [\[CrossRef\]](#)
49. Kitchenham, B.; Pearl Brereton, O.; Budgen, D.; Turner, M.; Bailey, J.; Linkman, S. Systematic literature reviews in software engineering—A systematic literature review. *Inf. Softw. Technol.* **2009**, *51*, 7–15. [\[CrossRef\]](#)
50. PRISMA: Preferred Reporting Items for Systematic Reviews and Meta-Analyses. Available online: <https://www.prismastatement.org> (accessed on 15 July 2025).
51. Guo, X.; Feng, Q.; Guo, F. CMTNet: A hybrid CNN-transformer network for UAV-based hyperspectral crop classification in precision agriculture. *Sci. Rep.* **2025**, *15*, 12383. [\[CrossRef\]](#) [\[PubMed\]](#)
52. Nyarko, F.; Tack, F.M.G.; Mouazen, A.M. Potential of visible and near infrared spectroscopy coupled with machine learning for predicting soil metal concentrations at the regional scale. *Sci. Total Environ.* **2022**, *841*, 156582. [\[CrossRef\]](#) [\[PubMed\]](#)
53. Shafique, A.; Siraj, M.; Cheng, B.; Alsaif, S.A.; Sadad, T. Hyperspectral imaging and advanced vision transformers for identifying pure and pesticide-coated apples. *IEEE Access* **2025**, *13*, 66405–66419. [\[CrossRef\]](#)
54. Sun, M.; Yang, Y.; Li, S.; Yin, D.; Zhong, G.; Cao, L. A study on hyperspectral soil total nitrogen inversion using a hybrid deep learning model CBiResNet-BiLSTM. *Chem. Biol. Technol. Agric.* **2024**, *11*, 157. [\[CrossRef\]](#)
55. Sun, Y.; Zhang, J.; Bai, J.; Xu, Y.; Chen, Y.; Lan, Y.; Han, X. Comprehensive assessment of soil quality in greenhouse agriculture based on genetic algorithm and neural network. *J. Soils Sediments* **2024**, *24*, 1302–1315. [\[CrossRef\]](#)
56. Xu, X.; Chen, S.; Ren, L.; Han, C.; Lv, D.; Zhang, Y.; Ai, F. Estimation of heavy metals in agricultural soils using Vis-NIR spectroscopy with fractional-order derivative and generalized regression neural network. *Remote Sens.* **2021**, *13*, 2718. [\[CrossRef\]](#)
57. Zhang, H.; Yin, A.; Yang, X.; Fan, M.; Shao, S.; Wu, J.; Wu, P.; Zhang, M.; Gao, C. Use of machine-learning and receptor models for prediction and source apportionment of heavy metals in coastal reclaimed soils. *Ecol. Indic.* **2021**, *122*, 107233. [\[CrossRef\]](#)
58. Guo, H.; Wang, J.; Zhang, D.; Cui, J.; Yuan, Y.; Bao, H.; Yang, M.; Guo, J.; Chen, F.; Zhou, W.; et al. Mapping surface soil organic carbon density of cultivated land using machine learning in Zhengzhou. *Environ. Geochem. Health* **2025**, *47*, 1. [\[CrossRef\]](#)
59. Raji, Z.; Ebtehaj, I.; Bonakdari, H.; Khalloufi, S. Artificial intelligence-driven assessment of critical inputs for lead adsorption by agro-food wastes in wastewater treatment. *Chemosphere* **2024**, *368*, 143801. [\[CrossRef\]](#)
60. Sahbeni, G. Soil salinity mapping using Landsat 8 OLI data and regression modeling in the Great Hungarian Plain. *SN Appl. Sci.* **2021**, *3*, 587. [\[CrossRef\]](#)
61. Tang, L.; Yin, Z.; Su, H.; Lyu, W.; Luo, B. WFSS: Weighted fusion of spectral transformer and spatial self-attention for robust hyperspectral image classification against adversarial attacks. *Vis. Intell.* **2024**, *2*, 5. [\[CrossRef\]](#)
62. Li, X.; Geng, T.; Shen, W.; Zhang, J.; Zhou, Y. Quantifying the influencing factors and multi-factor interactions affecting cadmium accumulation in limestone-derived agricultural soil using random forest (RF) approach. *Ecotoxicol. Environ. Saf.* **2021**, *209*, 111773. [\[CrossRef\]](#)
63. Patel, A.K.; Ghosh, J.K.; Pande, S.; Sayyad, S.U. Deep-learning-based approach for estimation of fractional abundance of nitrogen in soil from hyperspectral data. *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.* **2020**, *13*, 6495–6511. [\[CrossRef\]](#)
64. Ram Talib, N.S.; Halmi, M.I.E.; Abd Ghani, S.S.; Zaidan, U.H.; Shukor, M.Y.A. Artificial neural networks (ANNs) and response surface methodology (RSM) approach for modelling the optimization of chromium (VI) reduction by newly isolated *Acinetobacter radioresistens* strain NS-MIE from agricultural soil. *Biomed. Res. Int.* **2019**, *2019*, 5785387. [\[CrossRef\]](#)

65. Tatsumi, K.; Usami, T. Plant-level prediction of potato yield using machine learning and unmanned aerial vehicle (UAV) multispectral imagery. *Discov. Appl. Sci.* **2024**, *6*, 649. [\[CrossRef\]](#)
66. Wang, C.; Liu, B.; Liu, L.; Zhu, Y.; Hou, J.; Liu, P.; Li, X. A review of deep learning used in the hyperspectral image analysis for agriculture. *Artif. Intell. Rev.* **2021**, *54*, 5205–5253. [\[CrossRef\]](#)
67. Wu, M.; Qi, C.; Derrible, S.; Choi, Y.; Fourie, A.; Ok, Y.S. Regional and global hotspots of arsenic contamination of topsoil identified by deep learning. *Commun. Earth Environ.* **2024**, *5*, 10. [\[CrossRef\]](#)
68. Fendrich, A.N.; Van Eynde, E.; Stasinopoulos, D.M.; Rigby, R.A.; Mezquita, F.Y.; Panagos, P. Modeling arsenic in European topsoils with a coupled semiparametric (GAMLSS-RF) model for censored data. *Environ. Int.* **2024**, *185*, 108544. [\[CrossRef\]](#) [\[PubMed\]](#)
69. Genova, G.; Della Chiesa, S.; Mimmo, T.; Borruso, L.; Cesco, S.; Tasser, E.; Matteazzi, A.; Niedrist, G. Copper and zinc as a window to past agricultural land-use. *J. Hazard. Mater.* **2022**, *424*, 126631. [\[CrossRef\]](#)
70. Moharram, M.A.; Sundaram, D.M. Land use and land cover classification with hyperspectral data: A comprehensive review of methods, challenges and future directions. *Neurocomputing* **2023**, *536*, 90–113. [\[CrossRef\]](#)
71. Salas, E.A.L.; Kumaran, S.S. Hyperspectral bare soil index (HBSI): Mapping soil using an ensemble of spectral indices in machine learning environment. *Land* **2023**, *12*, 1375. [\[CrossRef\]](#)
72. Wei, H.; Huang, Y.; Santiago, P.J.; Labachyan, K.E.; Ronaghi, S.; Banda Magana, M.P.; Huang, Y.-H.; Jiang, S.C.; Hochbaum, A.I.; Ragan, R. Decoding the metabolic response of *Escherichia coli* for sensing trace heavy metals in water. *Proc. Natl. Acad. Sci. USA* **2023**, *120*, 2017. [\[CrossRef\]](#)
73. Zhang, Z.; Huang, L.; Wang, Q.; Jiang, L.; Qi, Y.; Wang, S.; Shen, T.; Tang, B.-H.; Gu, Y. UAV hyperspectral remote sensing image classification: A systematic review. *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.* **2025**, *18*, 3099–3124. [\[CrossRef\]](#)
74. Flores Peña, P.; Ale Isaac, M.S.; Gifu, D.; Pechlivani, E.M.; Ragab, A.R. Unmanned aerial vehicle-based hyperspectral imaging and soil texture mapping with robust AI algorithms. *Drones* **2025**, *9*, 129. [\[CrossRef\]](#)
75. Wang, Y.; Ma, H.; Wang, J.; Liu, L.; Pietikäinen, M.; Zhang, Z.; Chen, X. Hyperspectral monitor of soil chromium contaminant based on deep learning network model in the Eastern Junggar coalfield. *Spectrochim. Acta A Mol. Biomol. Spectrosc.* **2021**, *257*, 119739. [\[CrossRef\]](#)
76. Khan, U.; Khan, M.K.; Latif, M.A.; Naveed, M.; Alam, M.M.; Khan, S.A.; Su'ud, M.M. A systematic literature review of machine learning and deep learning approaches for spectral image classification in agricultural applications using aerial photography. *Comput. Mater. Continua* **2024**, *78*, 2967–3000. [\[CrossRef\]](#)
77. Ullah, H.; Khan, S.; Chen, B.; Shahab, A.; Riaz, L.; Lun, L.; Wu, N. Machine learning approach to predict adsorption capacity of Fe-modified biochar for selenium. *Carbon Res.* **2023**, *2*, 29. [\[CrossRef\]](#)
78. Xu, X.; Xu, Z.; Liang, L.; Han, J.; Wu, G.; Lu, Q.; Liu, L.; Li, P.; Han, Q.; Wang, L.; et al. Risk hotspots and influencing factors identification of heavy metal(loid)s in agricultural soils using spatial bivariate analysis and random forest. *Sci. Total Environ.* **2024**, *954*, 176359. [\[CrossRef\]](#)
79. Polechońska, L.; Klink, A.; Dambiec, M. Trace element accumulation in *Salvinia natans* from areas of various land use types. *Environ. Sci. Pollut. Res.* **2019**, *26*, 30242–30251. [\[CrossRef\]](#)
80. Neri, I.; Caponi, S.; Bonacci, F.; Clementi, G.; Cottone, F.; Gammaitoni, L.; Figorilli, S.; Ortenzi, L.; Aisa, S.; Pallottino, F.; et al. Real-time AI-assisted push-broom hyperspectral system for precision agriculture. *Sensors* **2024**, *24*, 344. [\[CrossRef\]](#)
81. Wang, J.; Norgaard, T.; Pugliese, L.; Carvalho, P.N.; Wu, S. Global meta-analysis and machine learning reveal the critical role of soil properties in influencing biochar-pesticide interactions. *Environ. Int.* **2024**, *193*, 109131. [\[CrossRef\]](#) [\[PubMed\]](#)
82. Pan, B.; Lei, J.; Pan, B.; Tian, H.; Huang, L. Dialogue between algorithms and soil: Machine learning unravels the mystery of phthalates pollution in soil. *J. Hazard. Mater.* **2025**, *482*, 136604. [\[CrossRef\]](#) [\[PubMed\]](#)
83. Bancheri, M.; Fusco, F.; Torre, D.D.; Terribile, F.; Manna, P.; Langella, G.; De Vita, P.; Allocca, V.; Loishandl-Weisz, H.; Hermann, T.; et al. The pesticide fate tool for groundwater vulnerability assessment within the geospatial decision support system LandSupport. *Sci. Total Environ.* **2022**, *807*, 150793. [\[CrossRef\]](#) [\[PubMed\]](#)
84. Sabour, M.R.; Sakhaie, P.; Sharifian, F. Trend analysis of machine learning application in the study of soil and sediment contamination. *Int. J. Environ. Sci. Technol.* **2024**, *21*, 8327–8336. [\[CrossRef\]](#)
85. Montalvo-Romero, N.; Montiel-Rosales, A.; Purroy-Vásquez, R.; Quechulpa-Pérez, P. Agro-technological systems in traditional agriculture assistance: A systematic review. *IEEE Access* **2023**, *11*, 123047–123069. [\[CrossRef\]](#)
86. Souza Silva, J.A.O.; Soares de Siqueira, V.S.; Mesquita, M.; Rodrigues Vale, L.S.; Bezerra da Silva, J.L.; da Silva, M.V.; Barcelos Lemos, J.P.; Nunes Lacerda, L.; Soranz Ferrarezi, R.; Fonseca Elias de Oliveira, H. Artificial intelligence applied to support agronomic decisions for the automatic aerial analysis images captured by UAV: A systematic review. *Agronomy* **2024**, *14*, 2697. [\[CrossRef\]](#)
87. Chen, X.; Yuan, F.; Ata-Ul-Karim, S.T.; Liu, X.; Tian, Y.; Zhu, Y.; Cao, W.; Cao, Q. A bibliometric analysis of research on remote sensing-based monitoring of soil organic matter conducted between 2003 and 2023. *Artif. Intell. Agric.* **2025**, *15*, 26–38. [\[CrossRef\]](#)

88. Alvarado, J.; Restrepo-Arias, J.F.; Velásquez, D.; Maiza, M. Disease detection on cocoa crops based on computer-vision techniques: A systematic literature review. *Agriculture* **2025**, *15*, 1032. [[CrossRef](#)]
89. Castro-Espichan, J.; Espinoza-Flores, J.; Montero, J.N.; Gamboa-Cruzado, J.; Oseda Gago, D.; Mendivel Gerónimo, R. Visión artificial con realidad aumentada para el desplazamiento de personas con discapacidad visual: Una revisión sistemática de la literatura. *Rev. Ibérica Sist. Tecnol. Inf.* **2021**, *E45*, 346–357.
90. Cárdenas-Quispe, A.; Vergaray-Mezarina, R.; Gamboa-Cruzado, J. Machine learning for malware detection on Android: Systematic review of the literature. *RISTI–Rev. Ibérica Sist. Tecnol. Inf.* **2021**, *45*, 318–331.
91. Khan, A.; Laghari, A.; Awan, S. Machine learning in computer vision: A review. *ICST Trans. Scalable Inf. Syst.* **2021**, *8*, 169418. [[CrossRef](#)]
92. Ram, B.G.; Oduor, P.; Igathinathane, C.; Howatt, K.; Sun, X. A systematic review of hyperspectral imaging in precision agriculture: Analysis of its current state and future prospects. *Comput. Electron. Agric.* **2024**, *222*, 109037. [[CrossRef](#)]
93. Sahu, P.; Singh, A.P.; Chug, A.; Singh, D. A systematic literature review of machine learning techniques deployed in agriculture: A case study of banana crop. *IEEE Access* **2022**, *10*, 87333–87360. [[CrossRef](#)]
94. Akbar, J.U.M.; Kamarulzaman, S.F.; Muzahid, A.J.M.; Rahman, M.A.; Uddin, M. A comprehensive review on deep learning assisted computer vision techniques for smart greenhouse agriculture. *IEEE Access* **2024**, *12*, 4485–4522. [[CrossRef](#)]
95. Albites-Tapia, A.; Gamboa-Cruzado, J.; Almeyda-Ortiz, J.; Moreno Lázaro, A. Chatbots for the detection of Covid-19: A systematic review of the literature. *Int. J. Adv. Comput. Sci. Appl.* **2022**, *13*, 991–998. [[CrossRef](#)]
96. Gamboa-Cruzado, J.; Carbajal-Jiménez, P.; Romero-Villón, M.; Mujica Ruiz, O.H.; Coveñas Lalupu, J.; Campos Miranda, M. Chatbots for customer service: A comprehensive systematic literature review. *J. Theor. Appl. Inf. Technol.* **2022**, *100*, 5587–5597.
97. Rejeb, A.; Abdollahi, A.; Rejeb, K.; Treiblmaier, H. Drones in agriculture: A review and bibliometric analysis. *Comput. Electron. Agric.* **2022**, *198*, 107017. [[CrossRef](#)]
98. Hasan, K.R.; Tuli, A.B.; Khan, M.A.-M.; Kee, S.-H.; Samad, M.A.; Nahid, A.-A. Deep-learning-based semantic segmentation for remote sensing: A bibliometric literature review. *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.* **2024**, *17*, 1390–1418. [[CrossRef](#)]

Disclaimer/Publisher’s Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.