Asian Journal of Water, Environment and Pollution. 2026;23(2):025460350.
doi: 10.36922/AJWEP025460350

REVIEW ARTICLE

Artificial intelligence-based approaches for Pb^{2+} detection in agricultural soils: A review

Zhenxing Wang[1], Haolei Gong[2], Kefan Yang[3]*, Wei Zou[4,5], and Hui Xiao[6]*

[1]School of Artificial Intelligence and Software Engineering, Hunan Geely Automobile Vocational and Technical College, Xiangtan, Hunan, China
[2]Sanuo Biosensing Co., Ltd., Changsha, Hunan, China
[3]School of Artificial Intelligence, University of Sussex, Ningbo, China
[4]School of Chemistry and Chemical Engineering, Southwest Petroleum University, Chengdu, Sichuan, China
[5]Nanchong Electronic Information Industry Technology Research Institute, Nanchong, Sichuan, China
[6]Department of Chemical Engineering, Faculty of Engineering, Universiti Malaya, Kuala Lumpur, Malaysia
(This article belongs to the *Special Issue: Frontiers in Sustainable Development of Ecology and Environment*)
***Corresponding authors:** Kefan Yang (ky252@sussex.ac.uk); Hui Xiao (s2178714@siswa-old.um.edu.my)

Received: November 12, 2025; Revised: December 22, 2025; Accepted: December 31, 2025; Published online: February 13, 2026

Abstract: Lead ion (Pb^{2+}) contamination in agricultural soils poses serious risks to ecosystem stability and human health due to its persistence and bioaccumulative toxicity. Traditional detection techniques are constrained by complex sample preparation, matrix interference, long analysis cycles, and limited *in situ* applicability. To overcome these limitations, artificial intelligence (AI)-based approaches have been increasingly adopted. Given the heterogeneity of agricultural soils and sensing signals, model selection is inherently scenario-specific: Support vector machine/support vector regression are suitable for complex matrices and multi-ion interference; artificial neural networks are preferred for low-concentration or multi-ion detection; convolutional neural networks are effective for high-dimensional, weak, or multi-modal signals; and least absolute shrinkage and selection operator/Ridge methods enable rapid, low-cost field screening. The performance of AI models depends on key factors—such as training dataset, hyperparameter optimization, and validation metrics. Coordinated optimization of these parameters enables robust, precise, and interpretable Pb^{2+} quantification. AI applications in soil Pb^{2+} detection are categorized into three main approaches: (i) single-modality approaches enhance sensitivity and specificity to address Pb^{2+} alloying and weak signal issues; (ii) multi-modal fusion strategies effectively mitigate interferences from complex soil matrices; and (iii) automated integrated platforms enable fast, field-deployable analyses while minimizing manual intervention. The approaches form a coherent technical chain that progressively addresses key bottlenecks in agricultural soil Pb^{2+} detection. Nevertheless, research in this area still faces challenges, including field adaptability, chemical speciation decoupling, and interference under variable soil conditions. Future research should focus on synergistic strategies that integrate materials, AI, and field applications. This review provides a targeted reference for the accurate tracking and control of Pb^{2+} contamination in farmland soils.

Keywords: Agricultural soils; Lead ion contamination; Artificial intelligence-based detection techniques; Meta-evaluation of artificial intelligence-based techniques; Applications of lead ion detection

1. Introduction

As a core dynamic natural entity on the Earth's surface that underpins crop growth, agricultural soil plays a fundamental role in sustaining plant development and ensuring food security by regulating nutrient cycling, water retention, and biological activity.[1] Its formation is attributed to the long-term integrated effects of multiple factors, including climate, biological processes, parent material, and topography.[2] Agricultural soil serves as a critical interface linking the farmland biosphere, atmosphere, and hydrosphere.[3] Its quality exerts a direct impact on the quality of farmland aquatic environments and surrounding atmospheric environments, thereby triggering cascading effects on the health of farmland ecosystems, the safety of crops, and the health of humans and animals.[4]

However, with the advancement of the Industrial Revolution and the Green Revolution, the scale and intensity of chemical production, extraction, application, and disposal have increased dramatically.[5] Against this backdrop, the sources of pollutants in agricultural soils have become increasingly diverse, encompassing both anthropogenic inputs and naturally occurring contaminants, thereby exacerbating the pollution burden on farmland soils.[6] Anthropogenic pollutants are spread through inadequate management of urban and hazardous wastes, the exploitation of minerals and raw materials, and their direct introduction through agricultural practices.[7] In addition, natural pollutants are also released in farmland areas, for instance, through the weathering of parent materials that liberate heavy metals, the enrichment of metals in regions with high geological background levels, or the deposition of volcanic ash and aeolian dust carrying elemental constituents.[8] The continuous accumulation of contaminants in agricultural soils has progressively exceeded their environmental carrying capacity, posing serious threats not only to crop growth but also to water, air, and biological communities.[9] At present, the complex and combined contamination of agricultural soils has emerged as a central issue in global environmental research.

Among the diverse pollutant systems entering agricultural soil environments, heavy metals—characterized by their strong stability, high accumulation potential, and bioavailability—have emerged as the predominant contaminant form.[10] They exert profound impacts on the structural integrity of agricultural soil ecosystems and the safety of agricultural products. In addition, heavy metals pose serious threats to human health.[11] Lead ion (Pb^{2+}) is a significant focus of research on heavy metal pollution due to its widespread presence and strong bioaccumulative toxicity.[12] The toxicity of Pb^{2+} to agricultural ecosystems exhibits pronounced specificity. In crops, Pb^{2+} in soils may inhibit chlorophyll synthesis and cell division, induce oxidative stress and reactive oxygen species production, reduce photosynthetic efficiency, and lower fruit set, with severe exposure potentially causing plant death.[13]

In humans, children absorb Pb^{2+} at a much higher rate (40–50%) compared to adults (3–10%).[14,15] After consuming lead-contaminated agricultural products, lead (Pb) accumulates in bones and impairs neural development.[16] In addition, Pb^{2+} exposure in adults likely results in liver and kidney damage, and inorganic Pb has been classified as a Group 2A probable carcinogen, posing long-term health risks to populations exposed through contaminated food.[14,15]

Efforts such as ensuring agricultural product safety, maintaining farmland ecosystem health, and supporting soil remediation require accurate, efficient, and spatially representative detection of Pb^{2+} in agricultural soils. Conventional methods for detecting heavy metal ions involve collecting soil samples through field drilling, which are then analyzed in laboratories using techniques such as inductively coupled plasma mass spectrometry (ICP-MS), atomic absorption spectroscopy (AAS), atomic emission spectroscopy, and X-ray fluorescence spectroscopy (XRF).[17] These methods offer high accuracy and sensitivity. However, although these techniques are applied to agricultural soils, they still face several limitations. Spatial variability across farmland plots and single-point sampling fail to capture the heterogeneous distribution of contaminants.[18] Furthermore, the sampling and analysis processes are time-consuming, labor-intensive, and costly, generate secondary waste, and cannot provide real-time guidance for field management.[19] Baydarashvili *et al*.[20] proposed a hydrated silicate-based method for soil heavy metal detection, employing color changes from interactions between heavy metal ions and silicate materials as indicators. While simple, cost-effective, and suitable for rapid screening, its sensitivity is insufficient for detecting the low Pb concentrations required for agricultural product safety.

Recent advances in artificial intelligence (AI) provide new opportunities to overcome these limitations. AI spans key areas such as natural language processing, computer vision, robotics, planning, and expert systems.[21] Through algorithms and models, these technologies enable the processing of multi-source

data, feature recognition, and trend prediction.[22] Their capabilities for efficiently handling large heterogeneous datasets and automatically identifying complex features make them highly suitable for agricultural soil detection. They generalize across different scenarios and provide real-time analysis, meeting the needs for in-field monitoring, accurate data, and visualization results.[23,24] By integrating sensor and interactive technologies, spectral and elemental data obtained in the field using portable devices, including XRF and near-infrared spectroscopy, are collected for analysis.[25] Machine learning algorithms (e.g., random forests [RFs] and neural networks) then process these data to achieve precise detection of key heavy metals.[25,26]

Despite these advances, AI-based Pb^{2+} detection in agricultural soils remains underdeveloped. Most existing AI studies focus on aqueous systems, while soil-specific challenges—including complex matrices, multi-ion interference, Pb^{2+} speciation effects, and limited model interpretability—are insufficiently addressed. Current research is fragmented, with a lack of systematic integration of Pb^{2+}-specific bottlenecks, interpretable AI frameworks, and standardized pathways for field deployment. Pb^{2+} is the key contaminant in agricultural soils and requires efficient and accurate detection methods. Developing such methods is crucial for protecting farmland ecosystems, ensuring the safety of agricultural products, safeguarding public health, and promoting sustainable agricultural development. This study systematically reviews the research progress, technical challenges, and future directions of AI-based Pb^{2+} detection in agricultural soils. It aims to provide both theoretical guidance and practical insights for advancing Pb^{2+} detection in agricultural soils, thereby supporting sustainable agricultural development.

2. Agricultural soil lead ion contamination

Agricultural soil, a foundational component that sustains food security and maintains the stability of agricultural ecosystems, plays a crucial role in determining the sustainability of agricultural production and the integrity of regional ecological systems. Yet, under the combined pressures of intensive farming, industrial emissions, and domestic waste accumulation, the issue of agricultural soil contamination has become increasingly pronounced. This challenge now stands as a major environmental concern that limits the progress of green agriculture and poses a significant threat to human well-being. Agricultural soil pollution typically encompasses heavy metal pollution, organic pollution, and emerging microplastic pollution. Among these, heavy metals—due to their chemical stability and bioaccumulation potential—constitute the most persistent threat to the soil-crop system.

Compared with common heavy metals such as cadmium (Cd) and zinc (Zn), Pb is more easily immobilized in soil with weaker mobility. Once accumulated, it is difficult to remove, and its residual period is much longer than that of most heavy metals.[27] Specifically, after entering the soil, the immobilization of Pb^{2+} is primarily achieved through chemical binding with organic matter, and is supplemented by carbonate precipitation and adsorption by hydrated metal oxides.[28] Due to its limited mobility and pronounced surface enrichment effect, coupled with an exceptionally long half-life of up to 2,000 years in soil, Pb^{2+} is regarded as a classic severe soil contaminant.[29] Furthermore, Pb^{2+} migrates and accumulates through the soil-plant system. Key links in its entry into the food chain from the environment include the passive diffusion and active uptake of Pb^{2+} by plant roots via Zn-iron (Fe) permease/natural resistance-associated macrophage protein transporters, which also serve as important pathways for human intake of Pb^{2+} through food.[30,31] Pb^{2+} pollution poses a long-term, multi-system, multi-organ threat to biological organisms and human health, and even low-level exposure may result in irreversible damage.

The sources of heavy metal contamination in agricultural soils are mainly classified into natural and anthropogenic categories.[32] Human activities are the dominant drivers of most heavy metal pollution, whereas natural contributions are relatively minor. This pattern has been consistently observed in monitoring studies of agricultural soils across multiple regions worldwide. Historically, large-scale agricultural activities driven by humans have employed substantial amounts of Pb^{2+}-based insecticides and modern agricultural chemicals containing Pb^{2+} impurities, resulting in the continuous accumulation of Pb in agricultural soils.[33]

In terms of fertilizers, phosphate fertilizers often contain Pb^{2+} and Cd^{2+} impurities originating from the phosphate rock used as the raw material.[34] The continuous application of these fertilizers progressively raises the concentrations of these heavy metals in the soil. Research has shown that Pb^{2+} levels in soils reach up to 1.9 times the regional background value, while Cd^{2+} concentrations exceed the agricultural soil risk screening threshold.[35,36] In addition, excessive application of micronutrient fertilizers enriched with Zn^{2+} and copper ions (Cu^{2+}) further elevates their accumulation in soil, aggravating heavy metal enrichment in localized areas.

Furthermore, industrial activities are one of the major driving factors of heavy metal pollution in agricultural soils. Pollutants from industries such as metallurgy, electroplating, and coatings contain substantial amounts of Pb.[35] The contribution of industrial sources to soil metal ions is illustrated in Figure 1. Quantitative assessments conducted via the positive matrix factorization model indicate that industrial sources account for approximately 59.60% of accumulated Pb^{2+} in soil, with Cd^{2+} and chromium ions (Cr^{3+}) contributing approximately 59.70% and 35.50%, respectively.[37] When industrial wastewater is used for irrigation of farmland without adequate treatment, it introduces heavy metals into the soil system. In many industrial areas across Europe, farmland adjacent to factories has been subject to long-term industrial discharge, causing markedly higher concentrations of Cd^{2+} and arsenic ions that exceed local agricultural soil quality standards and give rise to regional heavy metal pollution hotspots.[38]

In addition, transportation activities also contribute to the entry of heavy metal pollutants into agricultural soils. Although leaded gasoline has been gradually banned in most regions worldwide, historical Pb^{2+} deposition from its combustion, Zn^{2+} released from tire wear, and Cu^{2+} from brake component abrasion still reach farmland near roads through atmospheric dry and wet deposition or surface runoff.[39] In soils, these metals largely persist in low-solubility forms—Pb often occurs as stable compounds (e.g., $PbCO_3$ and $Pb_3[PO^4_2)$, and Zn^{2+} associates with clay particles to form relatively immobile complexes.[40] Monitoring of farmland adjacent to roads at site 31E shows that surface soil concentrations of Pb^{2+} and Zn^{2+} are 2.3- and 1.8-fold higher than in fields farther away, respectively, with contamination decreasing progressively with distance from the road.[36,41]

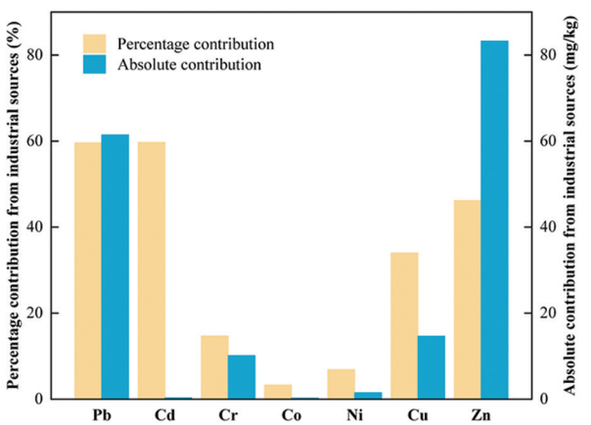

Figure 1. Percentage and absolute contributions of industrial sources to soil metal ions.[37] Data compiled from Ref.[37] and illustrated by the authors.

Abbreviations: Cd: Cadmium; Co: Cobalt; Cr: Chromium; Cu: Copper; Ni: Nickel; Pb: Lead; Zn: Zinc

Overall, agricultural soil constitutes the foundation of food security and the core hub of material cycling within farmland ecosystems, integrating the roles of production support, ecological regulation, and strategic resource storage. Its quality is fundamentally connected to the safety of agricultural products and human health. Currently, agricultural soils are threatened by Pb^{2+} contamination, which, due to its persistence and strong bioaccumulation potential, poses the most enduring risk to the soil-crop system. Pb^{2+} has become a major focus in studies of agricultural soil pollution because of its widespread occurrence, high toxicity, long half-life, and strong retention in soil environments. Therefore, this review primarily focuses on Pb^{2+}.

3. Agricultural soil lead ion detection technologies

3.1. Traditional methods

Heavy metal detection in agricultural soils is typically organized around the core framework of "sample pretreatment–separation-detection." The central objective is to realize efficient isolation of metals from complex soil matrices while maintaining precise qualitative and quantitative determination, thus fulfilling the demands of trace analysis and adaptation to heterogeneous substrates.[42]

During the pretreatment phase, research and technological development focus on removing interfering substances and effectively concentrating metals. Density separation, oil-assisted extraction, and electrostatic separation techniques are the major approaches. Hot plate digestion represents a classical approach, utilizing strong acids (e.g., nitric acid, hydrofluoric acid, and perchloric acid) with continuous heating over several hours to decompose the matrix.[43] However, this method consumes large volumes of solvents and may cause volatilization losses of thermally unstable metals. Microwave digestion enhances mass transfer through the application of microwave energy, completing digestion within 5–30 min while reducing solvent use to 10–30 mL.[44,45] Recoveries for elements like Pb^{2+} and Cr^{3+} reach 88–98%, substantially outperforming traditional hot plate digestion. Water bath digestion is widely applied to lightly contaminated soils for extracting easily soluble metals.[45,46] It is simple to operate but exhibits lower digestion efficiency.

In addition, sequential extraction methods—including the classic Tessier five-step procedure and the Community bureau of reference (BCR) three-step protocol—allow heavy metals to be classified into distinct chemical forms (e.g., exchangeable, carbonate-bound, and Fe–manganese oxide-bound), providing essential data for evaluating the bioavailability of elements like Pb. Nevertheless, the reagents used in these procedures lack strict selectivity, which may trigger redistribution of metals within the soil matrix. The degree of such interference varies among different metals, contributing to potential errors in speciation analysis.

In the analytical phase, the pursuit of high sensitivity and selectivity is paramount. The predominant detection techniques are generally divided into spectroscopic and electrochemical categories. Within the spectroscopic group, ICP-MS stands out for its exceptional sensitivity, achieving detection limits as low as 10^{-15}–10^{-12} g/kg.[47,48] It allows for simultaneous multi-element determination and isotopic characterization, which makes it especially effective for ultra-trace analysis of heavy metals such as Cd and mercury.[49,50] However, its practical use is limited by the considerable cost of the equipment and the need for operators with advanced technical expertise.

Inductively coupled plasma atomic emission spectrometry offers a broad linear dynamic range, typically spanning 4–6 orders of magnitude, and is well adapted for simultaneous determination of multiple elements at medium to high concentrations.[51,52] However, its analytical performance is hindered by spectral interferences. In comparison, AAS, encompassing both flame and graphite furnace modes, demonstrates strong selectivity.[53] Graphite furnace AAS reaches detection limits in the range of 10^{-12} to 10^{-9} g/kg, which renders it well suited for trace detection of elements like Pb and Cd.[54] Its main limitation, however, lies in the fact that only one element is measured at a time. XRF eliminates the need for sample digestion, allowing *in situ* screening to be completed within seconds to minutes, and facilitates rapid preliminary assessments of contamination hotspots.[55] Nevertheless, its analytical reliability is greatly influenced by matrix effects involving soil organic matter and moisture content, and its response to light elements remains comparatively low.

Electrochemical techniques, such as anodic stripping voltammetry, employ a preconcentration step to lower detection limits to the range of 10^{-9}–10^{-6} g/kg.[52,56] The instrumentation is portable and cost-effective, making it well-suited for on-site rapid measurements. However, the method is susceptible to interference from coexisting ions.

In practical soil heavy metal analysis, the detection of Pb^{2+} poses greater challenges compared with other elements, forming a key bottleneck that constrains both accuracy and stability. This issue manifests across three specific dimensions. On the one hand, Pb^{2+} loss and speciation interference during the pretreatment stage are difficult to control. Due to its high volatility, Pb tends to be lost with acid vapors during hot plate digestion at temperatures above 180°C, resulting in recoveries of only 78–85%.[57,58] Even when microwave digestion is employed, variations in soil organic matter content may lead to fluctuations in Pb recovery of 5–8%, considerably higher than the variability observed for Cr^{3+} (2–3%) and Cu^{2+} (3–5%).[45] In sequential extraction procedures, including the BCR method, Pb^{2+} exhibits even more pronounced redistribution. For instance, the acetate buffer (pH 5.0) simultaneously dissolves 10–15% of both organically bound and carbonate-bound Pb^{2+}, producing deviations in speciation analysis.[54,59] This deviation has a far greater impact on the assessment of Pb^{2+} bioavailability than it does for elements. On the other hand, matrix interferences during instrumental analysis of Pb^{2+} are especially difficult to eliminate. In inductively coupled plasma atomic emission spectrometry measurements, the Fe 238.20 nm line overlaps with the Pb^{2+} 238.32 nm line.[52] Given that soil Fe content commonly exceeds 20 g kg^{-1}, Pb^{2+} readings are consequently overestimated by 15–25%, a bias that conventional spectral subtraction techniques fail to fully correct.[52]

In XRF analysis, soil organic matter strongly influences the Pb^{2+} Lα characteristic line (10.55 keV).[60] For every 10% increase in organic matter content, the relative error (RE) in Pb^{2+} measurement rises by 3–5%. In agricultural soils with Pb^{2+} levels below 10 mg/kg, this error likely expands to −20% to 10%, whereas Cr^{3+} is less influenced by organic matter, exhibiting only −5% to 5% error.[48] During AAS detection, aluminum tends to form thermally stable alloys with Pb^{2+}, reducing atomization efficiency. The addition of a releasing agent is necessary, but it may introduce background absorption interference for Cd^{2+}.

Trace-level detection of Pb^{2+} shows poor consistency. For example, measurements of Pb^{2+} in the same agricultural soil by XRF and ICP-MS differ by approximately 24%, with an *F*-test confirming a significant difference ($p<0.05$).[48] By comparison, inter-method discrepancies for Cr^{3+} and Cu^{2+} are only 5–10%.[48,54] Moreover, variations in pretreatment parameters (e.g., digestion temperature, acid volume) and instrument settings among different operators result

in a relative standard deviation (RSD) of 5–8% for Pb^{2+} measurements, exceeding the variability observed for other heavy metals.[61]

Traditional methods for detecting Pb^{2+} in agricultural soils face significant challenges, being complex, time-consuming, and prone to partial Pb^{2+} loss. Moreover, Pb^{2+} may undergo speciation redistribution during pre-treatment and within complex soils, leading to poor consistency and limited precision in trace-level measurements. Addressing these limitations through optimized pre-treatment, improved analytical methods, and robust detection technologies is essential for accurate Pb^{2+} risk assessment in agricultural soils.

3.2. AI-based soil contamination detection technologies

3.2.1. AI-based technologies

AI-based technologies have emerged as a fundamental tool in soil science, providing data-driven modeling approaches that underpin precise and multidimensional soil management. Their applications primarily concentrate on two major areas: (i) prediction and evaluation of soil properties and (ii) management and dynamic monitoring of soil systems. The effectiveness of AI-based technologies strongly depends on appropriate model selection, as different algorithms exhibit distinct strengths under specific detection scenarios.

Regarding soil property prediction and evaluation, artificial neural networks (ANNs) integrate readily measurable parameters to estimate indicators that are difficult to obtain quickly, such as hydraulic conductivity and field water-holding capacity.[62] When combined with digital terrain models, ANNs promote the spatial accuracy of soil enzyme activity mapping.[63] In contrast, convolutional neural networks (CNNs) utilize visible and near-infrared spectral data to simultaneously estimate soil organic carbon, resulting in a reduction of approximately 30% in prediction error for deep soil carbon compared to traditional approaches.[64] CNNs enhance the accuracy of soil organic carbon detection through spectral data. RF and fuzzy logic systems assist in soil quality classification and preliminary pollution risk assessment. The RF model classifies soil quality based on indicators such as pH and organic matter, performing better than support vector machines (SVMs), while fuzzy logic systems are capable of identifying contaminated soils and evaluating risks using only a limited set of parameters.

In the field of soil management and dynamic monitoring, management-oriented modeling integrated with AI optimizes nitrogen fertilizer application to mitigate leaching, whereas AI-based analysis of soil-mineral interfacial interactions supports carbon sequestration. Portable Internet-of-Things-integrated AI systems (e.g., Agropad) enable the rapid assessment of key soil parameters, and soil digital twin technologies are capable of simulating dynamic soil processes.[65] Collectively, these advancements drive the modernization of agricultural soil management. Preventing and controlling soil pollution represents a critical component in safeguarding agricultural soil health. AI-based soil pollution detection technologies, capable of overcoming the limitations of traditional methods—including low efficiency and limited resistance to interference—have emerged as a major focus in contemporary soil science due to their ability to accurately identify pollutant types and concentrations. Nevertheless, realizing rapid and precise detection of soil contamination remains a major bottleneck.

For soil heavy metal detection, AI refines the core workflow of "signal analysis–interference control-concentration prediction," thereby addressing the constraints of traditional methods in accuracy, efficiency, and generalizability.[66] This is particularly evident for Pb^{2+}, where targeted technological optimizations have substantially strengthened detection performance. In spectral and electrochemical signal processing, CNNs combined with convolutional autoencoders execute dimensionality reduction and feature extraction on high-dimensional visible-near-infrared spectroscopy data. This effectively removes interference from soil organic matter and moisture, providing high estimation accuracy for Pb^{2+}, arsenic ions, and Cu^{2+} in mining soils. The models reach coefficients of determination (R^2) of 0.82–0.86. Specifically, the Pb^{2+}-focused model achieves an R^2 of 0.82, explaining 82% of the variability in Pb^{2+} concentration, and yields a root mean square error (RMSE) of only 0.63 ppm, indicating minimal random error in concentration prediction.[67] This performance markedly surpasses that of traditional ANNs and RF regression.

Signal overlap of Pb^{2+}, Cu^{2+}, and Zn^{2+} often interferes with Cd^{2+} detection, as Pb^{2+} signals are easily masked, causing quantitative deviations. SVM and RF models facilitate the disentanglement of non-linear relationships between current signals and interferences, which markedly increases the recovery of exchangeable heavy metals to 94.1–99.8%.[68] Pb^{2+} recovery reaches 98.8%, approaching the ideal value, and the corresponding concentration shows minimal deviation, ensuring accurate quantitative analysis.

In the concentration prediction and model optimization phase, the extremely randomized trees

(ExtraTrees) model integrates exchangeable heavy metal concentrations, soil properties (e.g., clay content, cation exchange capacity), and elemental descriptors to predict total heavy metal concentrations in soils at a global scale.[69,70] The model demonstrates high accuracy on the test set, with an R^2 of 0.90, explaining 90% of the variability in total soil heavy metal concentrations.[71] The predictions for typical heavy metals, including Pb^{2+}, show excellent fitting performance. SHapley Additive exPlanations (SHAP) values and partial dependence plots highlight exchangeable concentrations and elemental magnetic moments as key driving factors. Among them, the exchangeable Pb^{2+} concentration contributes most linearly to total Pb^{2+} levels, providing essential data support for tracing Pb^{2+} pollution sources. Meanwhile, Bayesian optimization efficiently selects model hyperparameters, reducing computational cost by over 60% compared with traditional grid search, and enhancing the model's generalization ability for Pb^{2+} prediction across different soil types.[71]

In summary, traditional methods, based on a "sample pretreatment-instrumental analysis" framework, have matured for heavy metals, yet they generally involve complex preparation, long analysis cycles, and susceptibility to matrix interference. Pb^{2+} detection remains challenging due to volatility and the overlap of multi-ion signals, highlighting a critical knowledge gap. AI approaches enhance signal interpretation and image analysis, optimizing efficiency and robustness. Notably, AI offers targeted solutions for the key difficulties in Pb^{2+} measurement, justifying the focus of this study on its application for precise and high-throughput soil Pb^{2+} monitoring. In preliminary Pb^{2+} detection, multiple AI models are typically employed to address the challenges posed by weak signals, high interference, complex soil matrices, and diverse field conditions.

3.2.2. Key parameters affecting lead ion detection

(a) Training data sets

The performance of soil Pb^{2+} detection models strongly depends on the extent to which the training dataset captures the complexity of the matrix.[72] For relatively simple sandy soils with low organic matter content (<3%), linear or weakly non-linear models, such as least absolute shrinkage and selection operator (Lasso) or Ridge (regularized linear regression variants), typically require no fewer than 35 effective samples to achieve stable predictive performance.[73,74] In contrast, for clay-rich soils with high organic matter (>5%), where organic complexation, clay adsorption, and multi-ion interference are more pronounced, non-linear models like SVM or CNN generally need at least 64 orthogonal experimental samples to ensure sufficient generalization.[75,76] The training dataset should systematically encompass variations in soil texture (e.g., sandy, clayey, and loamy), organic matter content (1–10%), and Pb^{2+} concentration ranges (0.1 nM–1,000 μM), enabling the models to adapt effectively to typical agricultural soil matrices.[75]

During dataset partitioning and model validation, the principle of consistent concentration distribution should be followed to avoid sample bias. Single-modality detection techniques, such as ANN, commonly adopt an 80:20 random split,[77] whereas multi-modal data, such as spectral–electrochemical fused CNN, are better suited to a 3:1 SPXY(sample set partitioning based on X–Y distances) partitioning method to preserve both feature–space and concentration–space consistency.[78] Furthermore, model validation should employ a dual-system approach combining laboratory standard samples and actual soil specimens. The soil samples should span a range of pH conditions (4.5–8.0) and typical interfering ion concentrations (Cu^{2+}/Zn^{2+}: 50–350 μg/L) to rigorously assess the model's robustness against soil-specific interferences.[75]

(b) Hyperparameter settings

Hyperparameter settings critically determine a model's capacity to resolve Pb^{2+} signals and resist matrix-induced interference. These settings should be tailored according to soil physicochemical properties, such as organic matter content and pH, as well as the type of sensing modality, including spectral, electrochemical, or visual signals.

In SVM/support vector regression (SVR) models, the radial basis function (RBF) kernel penalty parameter C typically ranges from 1 to 50. For soil spectral signals with low noise, smaller values (1–5) of C are preferred, while electrochemical signals subject to significant multi-ion interference require larger C values (20–30) to enhance model fitting.[75,76] The kernel coefficient γ is generally set between 0.01 and 0.5. High-dimensional spectral data often adopt $\gamma = 1/(\text{feature dimension} \times \text{variance})$ to prevent overfitting.[79] However, low-dimensional peak area features typically use 0.3–0.5 to improve signal discrimination (064). Parameter optimization methods include grid search, particle swarm optimization (PSO), and Bayesian optimization. Grid search is suitable for low-dimensional

parameter spaces, PSO performs well for non-linear, multi-interference scenarios, and Bayesian optimization effectively reduces computational cost while maintaining high predictive accuracy in high-dimensional feature settings.

Key hyperparameters of ANN (e.g., multilayer perceptron [MLP]) models include the number of hidden layer nodes (4–20), learning rate (0.001–0.1), and regularization strength (0.0001–0.01).[80] Scenarios involving multi-ion detection or signal superposition typically require more hidden nodes (8–20), whereas weak spectral signals benefit from lower learning rates to improve fitting accuracy. In soils with high organic matter or salinity, applying L2 regularization and selecting λ through cross-validation can effectively mitigate overfitting caused by matrix-induced noise.[81]

Convolutional neural network model parameters balance feature extraction and hardware constraints. Convolutional kernels are typically 3 × 3 with 32–64 filters, with larger sizes used for complex matrices.[82] Dropout rates are set at 0.2–0.3 to reduce overfitting, batch sizes are typically 16–32, and training epochs may reach 80–100 for low-concentration Pb^{2+} detection.[74] Common optimization strategies include random search, early stopping, and model pruning. Pruning effectively reduces model size while maintaining accuracy, which enhances field deployment efficiency.

In addition, Lasso/Ridge regression typically uses $\alpha = 0.001–0.1$ to remove scattered or background interference features.[83] Conversely, Ridge regression often adopts $\alpha = 0.5–1.0$ for visual or colorimetric signals to mitigate signal drift induced by salinity.[84] In soils with high organic matter or salinity, regularization strength (C, λ, or α) should be appropriately increased to suppress matrix-induced noise. Alternatively, for low-concentration Pb^{2+} detection (<1 nM), a lower learning rate and extended training epochs are required to enhance the model's sensitivity to weak signals.

(c) Validation metrics

Model performance evaluation is recommended to address three dimensions: accuracy, anti-interference capability, and practicality. Accuracy is typically assessed using $R^2 \geq 0.93$ and RMSE $\leq$ 15 μg/L.[75,85] Anti-interference capability is quantified by RE $\leq$ 6.5% and RSD $\leq$ 8%.[86,87] Practicality focuses on recovery rate (95–105%) and single-sample detection time (≤75 min).[88] All measurements are benchmarked against ICP-MS, AAS, or XRF, and deviations in actual soil samples are controlled within 5% to ensure reliability and traceability for agricultural pollution assessment.[75]

As a whole, Pb^{2+} detection performance in agricultural soils arises from the interplay of training dataset design, hyperparameter tuning, and validation strategy rather than any single model. Comprehensive training samples capturing soil matrix variability and interference are essential for stable generalization. Hyperparameters aligned with soil properties and sensor signals govern the model's ability to resolve multi-ion interference and weak Pb^{2+} signals. Standardized, multidimensional performance metrics ensure comparability and transferability in field applications. Coordinated optimization of these factors lays a solid foundation for translating AI models from laboratory validation to reliable field deployment.

3.2.3. The interpretability of AI-based models

In soil Pb^{2+} detection, existing AI models exhibit three major categories of interpretability challenges. First, the characterization of interference sources remains ambiguous. For example, Liu *et al*.[75] reported that a feature-SVR model selected numerous Pb^{2+}-related features, yet the high contribution of specific potential intervals lacked mechanistic justification, and the temporal effects of the interfering ion Cu^{2+} were not explicitly represented. Second, the linkage to chemical speciation is insufficient. In the ANN model developed by Kudr *et al*.,[77] prediction accuracy declined in soils with high organic matter content, failing to capture the influence of organic binding on the bioavailable Pb^{2+} fraction. Third, field adaptability is questionable. The Lasso regression proposed by Yu *et al*.[74] showed pronounced performance variability across soil types, while explanations for salinity interference or for adjustments of regularization parameters were largely absent.

In response to these issues, prior studies have explored interpretability from the perspectives of features, decision processes, and chemical mechanisms. At the feature level, methods such as SHAP are employed to quantify the contributions of input variables. For instance, ANN-based models are able to identify Cu^{2+} interference thresholds and rationalize the selection of Pb^{2+} stripping peaks. At the decision level, tools including local interpretable model-agnostic explanations are used to construct local linear approximations, linking abrupt variations in fluorescence or colorimetric signals to changes in Pb^{2+} concentration. Coupling with chemical mechanisms

further integrates stripping thermodynamic parameters or alloy adsorption principles into model architectures, aligning predictions closely with electrochemical theory, which enhances their verifiability.

Based on the analysis of interpretability limitations in existing AI models for soil Pb^{2+} detection, four optimization directions are proposed to enhance interpretability:

(i) Hybrid models with embedded physical mechanisms: Incorporate electrochemical or spectroscopic parameters as constraint layers within AI architectures, thereby endowing model parameters with clear physical meaning while reducing reliance on labeled data.

(ii) Hierarchical SHAP for interference tracing: Develop a layered SHAP framework to quantify and accurately trace the contributions of matrix and ionic interferences.

(iii) Region-adaptive meta-learning framework: Combine meta-learning with local interpretability techniques to enable rapid adaptation across soils from different regions and generate site-specific explanation reports.

(iv) Chemistry-signal mapping database: Establish a reference linking Pb^{2+} chemical forms to detection signals, allowing models to output both concentration and speciation proportions, supporting pollution risk assessment.

In summary, by integrating feature contribution analysis, decision visualization, chemical mechanism embedding, and database support, AI models can transition from black-box predictions to a mechanism-driven, multidimensional interpretability framework. This approach enables traceable interference sources, improved regional adaptability, and explicit links to chemical speciation, providing a robust foundation for precise detection and field-level application.

4. AI-based techniques: Meta-evaluation for lead ion detection in agricultural soils

4.1. Evaluation of the relationship between AI-based techniques and lead ion detection

Agricultural soils present several critical challenges for Pb^{2+} detection. These include complex matrices with coexisting clay minerals and organic matter, significant multi-ion interference, and the need for high adaptability in field measurements. To tackle these issues, various AI techniques have been applied, exhibiting notable performance advantages. The characteristics and suitability of different approaches are summarized as follows:

(a) Support vector machine and support vector regression

Support vector machine and its regression variant (SVR) employ kernels to map low-dimensional sensor signals into high-dimensional feature spaces. This enables non-linear fitting of Pb^{2+} signals while effectively handling multi-ion interference and matrix-induced noise. Baseline correction and feature selection help mitigate interference from organic fluorescence and clay scattering. Studies indicate that, for Cu^{2+}/Pb^{2+} interference in high-organic soils, Feature-SVR reduces detection error by 83% and achieves a recovery rate of 107.10%.[75] RBF–SVM models across varying pH conditions improve accuracy by approximately 27% in red and alluvial soils.[76] Rapid parameter optimization allows SVR to meet real-time requirements for automated devices, shortening single-sample detection time from 45 s to 12 s while maintaining $R^2 \approx 0.975$.[75,89] Overall, SVM/SVR exhibit strong adaptability to complex matrices and multi-ion interference, with feature selection and parameter tuning being key factors for performance enhancement.

(b) Artificial neural networks

Multilayer perceptrons separate Pb^{2+} signals from other ions through non-linear transformations in hidden layers, while regularization strategies suppress overfitting caused by matrix noise. In scenarios with multi-ion coexistence or cross-reactive sensor arrays, ANNs achieve high-precision detection with $R^2 > 0.99$, maintaining sensitivity to low-concentration signals.[77,80] Their multilayer topology and strong signal decoupling capability make ANNs particularly effective for detecting low-concentration Pb^{2+} amid multiple interfering ions. However, training duration and parameter selection markedly influence model performance.

(c) Convolutional neural networks

One-dimensional convolutional architectures automatically extract deep features from high-dimensional sensor signals, efficiently handling weak Pb^{2+} signals in surface-enhanced Raman spectroscopy (SERS) spectra or multi-modal data. Dropout layers and model pruning suppress overfitting and reduce model size, facilitating deployment on portable detection devices. Experiments demonstrate that CNNs achieve detection limits as low as 0.4 nM in sandy and high-organic soils, improve resistance to scattering interference by 40%, and substantially accelerate detection.[74,78] When CNNs excel with high-dimensional, weak, and multi-modal signals, they require relatively high computational resources.

(d) Regularized linear regression variants

The least absolute shrinkage and selection operator induces sparsity in feature weights, thereby removing soil matrix-related interference features. By comparison, Ridge constrains the squared magnitude of coefficients to improve model stability under signal fluctuations. Both methods exhibit low computational complexity and are readily deployable on smart terminals, enabling rapid and low-cost in-field detection. In high-salinity soils or during *in situ* sampling, Lasso and Ridge achieve R^2 values above 0.99 and markedly reduce detection errors.[73] These models are particularly suitable for cost-sensitive and rapid field screening, with performance being most strongly influenced by feature selection and regularization strength.

Taken together, AI techniques show differentiated suitability for Pb^{2+} detection in agricultural soils. SVM/SVR excel under complex matrices and multi-ion interference. ANN is beneficial for multi-ion coexistence and low Pb^{2+} levels. CNN is advantageous for high-dimensional weak and multi-modal signals, while Lasso/Ridge support low-cost and rapid field screening. Across methods, functional feature preprocessing and targeted parameter optimization are essential for soil adaptability, enabling reduced detection errors, improved recovery, and practical field deployment.

4.2. A meta-evaluation of soil lead ion detection performance

Based on a systematic review and comparative analysis of existing studies, the key performance metrics of various AI models for Pb^{2+} detection in agricultural soils are summarized in Table 1. The meta-analysis reveals notable differences among models. Although all optimize for Pb^{2+}-specific response features, they differ in detection sensitivity, interference resistance, operational efficiency, and engineering feasibility. These differences highlight each model's practical limits and trade-offs in field applications.

Based on Table 1, RBF-SVM performs robustly under complex matrices and multi-ion interference, but declines at ultra-low Pb^{2+} concentrations and high-salinity conditions. ANN (MLP) enables accurate topsoil screening with strong signal decoupling, yet relies on large training sets. CNN excels in high-organic soils with enhanced weak-signal detection, though it requires high-cost computational resources. Lasso/Ridge regression offers fast, low-cost field screening but loses accuracy under strong interference. Hybrid Wavelet-Support Vector Regression (H-W-SVR) balances sensitivity and engineering feasibility, meeting typical soil Pb^{2+} background levels, though cross-soil generalization needs improvement.

In summary, the strengths of AI models for Pb^{2+} detection in agricultural soils lie not in single metrics, but in their adaptation to specific soil types, interference levels, and use scenarios. Complex models excel in high-precision analysis of low-concentration Pb^{2+} in organic-rich soils, while lightweight models suit rapid, low-cost field screening. Future Pb^{2+} detection should move from individual model optimization toward integrated "model-soil-engineering" design to simultaneously enhance sensitivity, robustness, and scalability.

4.3. Technical bottlenecks in lead ion detection and AI-driven breakthrough

The environmental chemistry of Pb^{2+}—characterized by strong alloying tendency, high affinity for organic matter, and weak-acid-dominated dissolution—creates

Table 1. Comparative performance of artificial intelligence-based models for lead ion detection in agricultural soils

Performance metrics	SVM (RBF)	ANN (MLP)	CNN	Lasso/ridge regression
Quantitative accuracy	LOD=5.7±4.8 Nm; R^2=0.94±0.04	LOD=6.3±5.5 nM; R^2=0.93±0.05	LOD=0.4±0.1 nM; R^2=0.98±0.01	LOD=0.05±0.01 nM; R^2=0.99±0.003
Interference resistance	RE=5.2±1.0%	RE=6.5±2.0%	RE=3.8±0.5%	RE=2.1±0.3%
Detection efficiency	25±10 min	15±7.5 min	35±15 min	10±5 min
Equipment cost (RMB)	50,000–150,000	<50,000	>200,000	<30,000
Soil scenario adaptability	Detection in complex matrices with multi-ion interference	Rapid screening in surface soils	High-precision detection in high-organic-matter soils	On-site, low-cost *in situ* detection

Note: Data adapted from Refs.[72,74,75,78]

Abbreviations: ANN: Artificial neural network; CNN: Convolutional neural network; Lasso: Least absolute shrinkage and selection operator; LOD: Limit of detection; MLP: Multilayer perception; RBF: Radial basis function; RE: Relative error.

critical bottlenecks for its detection. To solve these problems, AI-driven approaches leverage targeted feature extraction and mechanism-informed constraints, enhancing signal detection and model generalization. Specific strategies include: extracting multidimensional features of alloying peaks to mitigate peak distortion; employing dual-mode signal fusion and regularization to improve recognition of weak signals; and incorporating regional soil physicochemical parameters to optimize model generalizability.

A systematic summary of these technical bottlenecks and corresponding AI solutions is presented in Table 2, facilitating intuitive comparison of different AI strategies and their effectiveness in Pb^{2+} detection.

5. Application of AI in lead ion detection in agricultural soils

Lead ion, as a representative non-biodegradable heavy metal, may trigger neurotoxicity, impair renal function, and accumulate within ecosystems even at sub-nanomolar concentrations. This underscores the urgent need for sensitive, on-site, and interference-resilient detection in environmental monitoring. Agricultural soil Pb^{2+} analysis faces three primary challenges: Complex matrices, interference from coexisting ions, and limited field applicability. Coupling AI with diverse sensing modalities—optical, electrochemical, and colorimetric or fluorescent—enables the precise interpretation of detection signals, systematic mitigation of confounding factors, and automated workflows. Current approaches are categorized into single-modality AI detection, multi-modal AI fusion, and automated integrated AI systems. These strategies span from foundational sensor optimization to advanced interference management and practical field deployment, forming a coherent technical chain that progressively addresses key bottlenecks in agricultural soil Pb^{2+} detection.

5.1. Single-modality AI-based detection technology

Single-modality technology forms the fundamental application of AI in soil Pb^{2+} detection. Characteristic Pb^{2+} signals are captured using a single sensing principle—such as optical, electrochemical, or colorimetric/fluorescent sensors analyzed via AI algorithms, including ANNs, SVM, and Lasso regression—which

Table 2. Artificial intelligence-based strategies for addressing key challenges in lead ion detection

Technical bottleneck	Pb^{2+}-specific manifestation	Artificial intelligence breakthrough path
Multi-ion interference (mainly Cu^{2+}/Zn^{2+})	Cu^{2+}forms Cu-Pb alloys, causing peak shifts (−0.54V Pb^{2+}peak overlaps with Cu^{2+}peak); high Zn^{2+}concentrations suppress Pb^{2+} stripping current	1. Peak-area feature mining (A_algo-SVR): Capture Pb^{2+}peak variations, RMSE reduced from 32.20 μg/L to 13.16 μg/L 2. 2D-COS analysis of interference sequences: Prioritize Cd^{2+}→Pb^{2+}→$Bi3^{+}$→Cu^{2+}correlated features
Strong matrix adsorption (organic matter/clay)	Pb^{2+}binding to humic acids (K=$10^{4.7}$) weakens signals (<5 μA at <1 nM); clay minerals induce SERS baseline drift	1. Dual-mode signal fusion (fluorescence+electrochemistry): CNN extracts 1,050 cm^{-1} Pb^{2+}peak, MAE=0.60 μg/L 2. Swab-type COF sensor+Lasso regression: Remove scattering redundancy, LOD=50 pM
Field signal fluctuation (pH/salinity/dissolved state)	Acidic red soils (pH <5.5) proton interference; saline-alkaline soils cause colorimetric signal drift ([G+B]/R±0.05)	1. Federated learning to update regional models: Optimize proton interference correction for southern red soils, RE reduced from 8.7% to 3.8% 2. H-W-SVR combining peak height+width: adapt to weakly acidic dissolved Pb^{2+}, LOD=0.33 μg/L
Decoupling of chemical speciation and measured values	Exchangeable Pb^{2+}15.9–53.2%, organic-bound 30–60%; models cannot differentiate, causing prediction bias >10%	1. Coupling BCR-extracted data: Use exchangeable/organic-bound fractions as constraints, ANN prediction versus speciation analysis R^2=0.98 2. LIME visualization: Interpret (G+B)/R inflection points in relation to Pb^{2+}-COF binding mechanism

Note: Adapted from Refs.[90,91]

Abbreviations: ANN: Artificial neural network; B: Blue; BCR: Community Bureau of Reference; COF: Covalent organic framework; COS: Correlation spectroscopy; G: Green; H-W-SVR: Hybrid Wavelet–Support Vector Regression; LIME: Local interpretable model-agnostic explanations; LOD: Limit of detection; MAE: Mean absolute error; R: Red; RE: Relative error; RMSE: Root mean square error; SERS: Surface-enhanced Raman spectroscopy; SVR: Support vector regression.

are applied to enhance signal processing and improve concentration prediction.[92] This approach mainly addresses the limitations of conventional single-modality detection, including poor specificity and susceptibility to single-source interferences. Considering the complex composition of agricultural soils, these techniques increase selective Pb^{2+} responses through algorithmic refinement or sensor material modification, providing a foundation for subsequent detection in more complex matrices.

5.1.1. Optical single-modality AI-based detection of lead ions in soils

Optical sensing techniques are widely employed in soil Pb^{2+} detection due to their rapid response and non-contact operation. They generate characteristic signals through optical interactions between Pb^{2+} and the sensing material.[93] AI algorithms are then applied to suppress interferences from soil organic matter fluorescence and particulate scattering, enabling accurate quantitative analysis of Pb^{2+}.

(a) Fiber Bragg grating sensor-ANNs coupling technology

The fiber Bragg grating (FBG)-ANN coupling approach was developed by Ghosh *et al*.[80] as a Pb^{2+} detection system based on a hybrid fiber grating configuration, combining long-period gratings (LPG) and FBG with nanocomposite materials (Figure 2). A feedforward ANN was introduced to calibrate the non-linear response, enabling accurate quantification of Pb^{2+}. In this design, the FBG simultaneously monitors ambient temperature to eliminate thermal cross-interference (Figure 2A). At the same time, the LPG serves as the Pb^{2+} detection unit, with its surface coated by an ammonia-doped graphene oxide-chitosan-polyacrylic acid nanocomposite (Figure 2B). Both elements are integrated in a single-mode optical fiber, and signal acquisition is achieved using a 785 nm laser source coupled with a spectrometer. The inset shows the transmission spectrum of the hybrid fiber grating, where the prominent peak corresponds to the LPG and the smaller peak to the FBG. The multilayer structure of the nanocomposite provides abundant chelation sites, including $–NH_2$ and –COOH, which specifically bind Pb^{2+} and induce a shift in the LPG cladding mode resonance wavelength. This design achieves a sensor sensitivity of 2.547 nm/nM and a limit of detection (LOD) as low as 0.5 nM, outperforming conventional tilted FBG sensors (LOD = 1.2 nM) and fiber surface plasmon resonance sensors, which, despite a LOD of 0.158 nM, exhibit a lower sensitivity of 2.101 nm/nM.

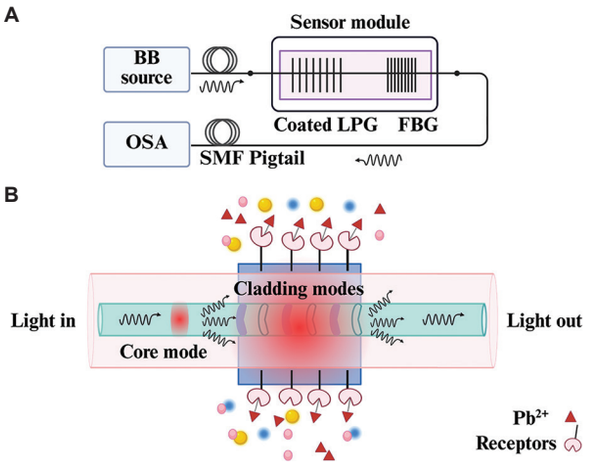

Figure 2. Schematic illustration of (A) the hybrid fiber grating sensor device and (B) the operating principle of the LPG. Image created by the authors.
Abbreviations: BB: Broadband light source; FBG: Fiber Bragg gating; LPG: Long-period grating; OSA: Optical spectrum analyzer; SMF: Single-mode fiber.

To mitigate non-linear responses caused by heterogeneous Pb^{2+} adsorption, a feedforward ANN with two hidden layers (20 nodes each) was trained using Pb^{2+} concentration and adsorption isotherm data as inputs, with LPG $LP_{0,9}$ wavelength shift ($\Delta\lambda_{0,9}$) as output.[94]

Figure 3 illustrates the architecture of the ANN. The predicted $\Delta\lambda_{0,9}$ values exhibit a high correlation with experimentally measured values, achieving a correlation coefficient (r) of 0.9999.[95] Within the 0–5 nM concentration range, the model demonstrates excellent linear response performance, with a linear fit R^2 value of 0.9969. This ANN model achieves outstanding fitting accuracy for the Toth adsorption isotherm, with an R^2 value as high as 0.9997. Its exceptional fitting capability ensures detection stability across the broad concentration range from 0.5 nM to 1,000 μM. As shown in Figure 3, the sensor displays a strong linear response in the low concentration range, where both the neural network prediction curve and the Toth-fitted curve closely match the experimental data. Moreover, the model precisely captures the sensor's non-linear response characteristics (quantified by $|\lambda FBG - \lambda LPG_{0,9}|$), demonstrating outstanding consistency with observed results.

Although initially designed for aqueous systems, the sensor is applicable to Pb^{2+} detection in agricultural

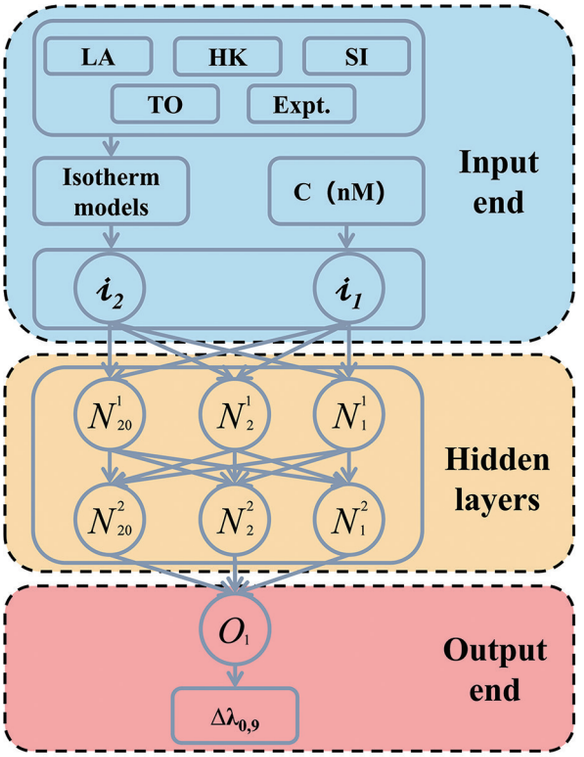

Figure 3. Schematic illustration of the artificial neural network architecture. Image created by the authors.

Abbreviations: C: Concentration; Expt.: Experiment; HK: Hidden layer kernel; LA: Learning algorithm; SI: Signal intensity; TO: Target output.

soil leachates by optimizing the nanocomposite to resist adsorption by soil clay minerals, thereby minimizing non-specific binding. The sensor's practicality was validated through spiked recovery experiments in tap water. For Pb^{2+} spiked samples at 0.5 nM, 5 nM, and 10 nM, the recoveries were 95.05%, 97.32%, and 98.72%, respectively, with REs ranging from 2.48% to 12.85%. In selectivity tests, the $\Delta\lambda_{0,9}$ response to 5 nM Pb^{2+} was significantly higher than for interfering ions of the same concentration, including Hg^{2+}, Cu^{2+}, and Cd^{2+}.

(b) Surface-enhanced Raman spectroscopy-machine learning integrated technology

Park *et al.*[76] constructed a Pb^{2+} detection system by integrating SERS with machine learning. The overall system architecture, sample preparation procedure, and signal acquisition workflow are clearly illustrated in Figure 4. The core experimental setup consists of a 785 nm laser source (Figure 4A), a commercial SERS substrate, and a spectrometer (Figure 4B). For sample preparation, 2.5 μL of $Pb(NO_3)_2$ solution was deposited onto the substrate and allowed to air-dry at room temperature. During detection, the laser power was maintained at 200 mW with a 500 ms exposure time, and signals were collected every 10 s to prevent degradation. They designed cross-batch benchmark datasets (Batch 1 and Batch 2), each containing 500 negative samples (Pb^{2+} concentration <0.01 μM) and 1,500 positive samples (Pb^{2+} concentration ≥ 0.01 μM), with positive samples covering concentrations of 0.01 μM, 0.1 μM, 10 μM, and 1,000 μM. This design simulates device and operational variability encountered in practical detection scenarios.[96]

To reduce the impact of background noise and batch effects on SERS signal detection accuracy, the study compared three preprocessing methods: Raw data (RAW), power spectrum normalization (PSN), and iterative constrained least squares baseline correction (BC). The analysis revealed that RAW data, containing substantial background noise, led to overlapping clusters across batches, whereas PSN excessively amplified noise, causing the model to misclassify all samples as positive. Conversely, BC effectively removed low-frequency background noise, producing clear clustering boundaries between positive and negative samples across batches, significantly reducing batch effects and lowering background noise intensity by 52% compared with raw data. The preprocessing effect on Pb $(NO_3)_2$ SERS spectra is depicted in Figure 5.

For machine learning model selection, an RBF-SVM was adopted, with ℓ_2 regularization penalty coefficient $C = 1$ and kernel coefficient $\gamma = 1/(2{,}000\times\sigma_{si}^2)$ (σ_{si}^2 representing spectral variance) to balance model fitting and generalization. Comparison with six other algorithms—logistic regression, linear kernel SVM, naïve Bayes, decision tree, RF, and MLP—demonstrated that RBF–SVM performed best for Pb^{2+} detection. Its cross-batch average balanced accuracy reached 84.6%, outperforming MLP (82.5%) and NB (78.9%). Notably, when Batch 2-trained models were applied to Batch 1 samples, a remarkable batch average balanced accuracy of 93.4% was achieved.[97]

5.1.2 Electrochemical single-modality AI-based detection of lead ions in soils

Electrochemical techniques, owing to their portability and low cost, constitute a key approach for *in situ* detection of Pb^{2+} in soils. These methods rely on the

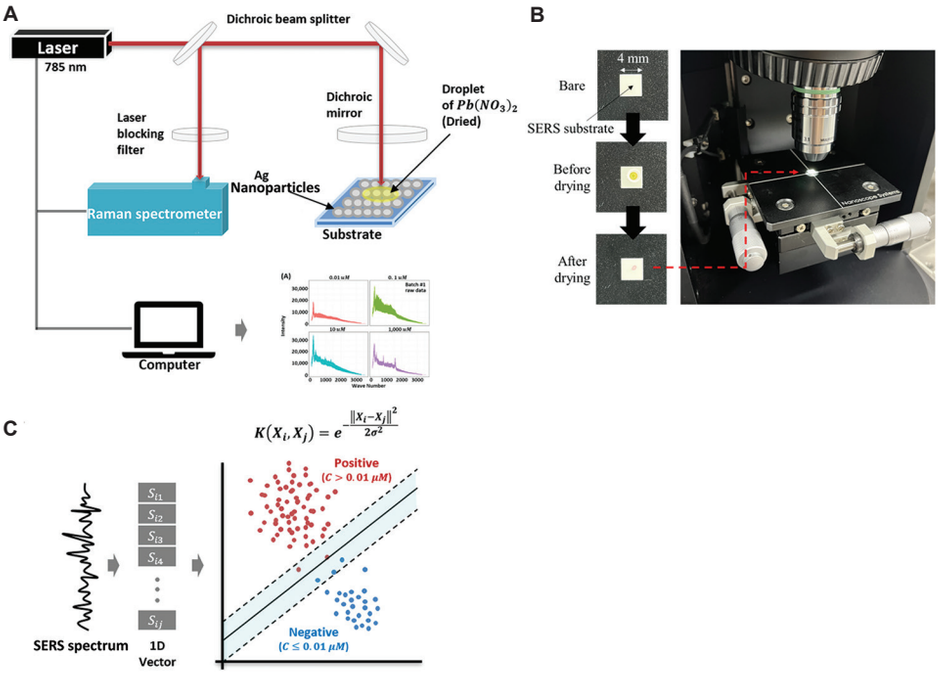

Figure 4. (A) Schematic illustration of the proposed SERS-based $Pb(NO_3)_2$ molecular detection system. (B) Sample preparation and signal acquisition steps. (C) Hypothetical decision boundary learned by the proposed RBF–SVM model. Adapted from Ref[76].

Abbreviations: 1D: One-dimensional; Ag: Silver; $Pb(NO_3)_2$: Lead(II) nitrate; RBF: Radial basis function; SERS: Surface-enhanced Raman spectroscopy; SVM: Support vector machine.

redox reactions of Pb^{2+} at the electrode surface to generate current or potential signals. By integrating AI algorithms, interferences from coexisting ions and soil organic matter are efficiently suppressed, improving detection accuracy.

(a) Thin-film mercury electrode-ANNs automated electrochemical detection

Kudr *et al*.[77] engineered an automated electrochemical detection system combining a thin-film mercury electrode (TFME) with an ANN for the simultaneous determination of Pb^{2+}, Zn^{2+}, Cd^{2+}, and Cu^{2+} in environmental samples. Its core architecture, as illustrated in Figure 6, comprises an electrochemical robotic platform (Figure 6A), a three-electrode system (Figure 6B), and ANN-based signal analysis (Figure 6C). This robotic platform employs motorized *x*-, *y*-, and *z*-axis controls to precisely position electrodes within a 24-well microplate. The three-electrode system comprises a TFME electrode with a carbon tip formed by electrodeposited mercury as the working electrode, a silver/silver chloride reference electrode, and a platinum wire counter electrode. An automated pipetting system permits the mixing of samples and buffer solutions. Relative to bare carbon electrodes, the TFME supplies a wide cathodic potential window and high adsorption capacity due to the mercury film, thus allowing the detection of metal ion signals that bare electrodes cannot capture and lowering the Pb^{2+} LOD from 0.2 μg/mL to 0.03 μg/mL.[98]

To mitigate signal interference in Pb^{2+} detection arising from the coexistence of multiple metals, initial attempts employed linear regression and linear-exponential hybrid numerical regression models. However, the R^2 value for Cu^{2+} concentration was only 0.87, which is insufficient to meet the accuracy requirements.[77] Consequently, an MLP type ANN model was constructed with a 4–8–4 architecture: four input neurons corresponding to the peak heights of the four metals, eight hidden neurons, and four output neurons representing the metal concentrations, as illustrated in Figure 7. An exponential activation function was applied in the hidden layer, and a logistic activation function was used in the output layer. After 1,237 training cycles, the model achieved R^2 values above 0.99 for the training, testing, and validation sets. The final

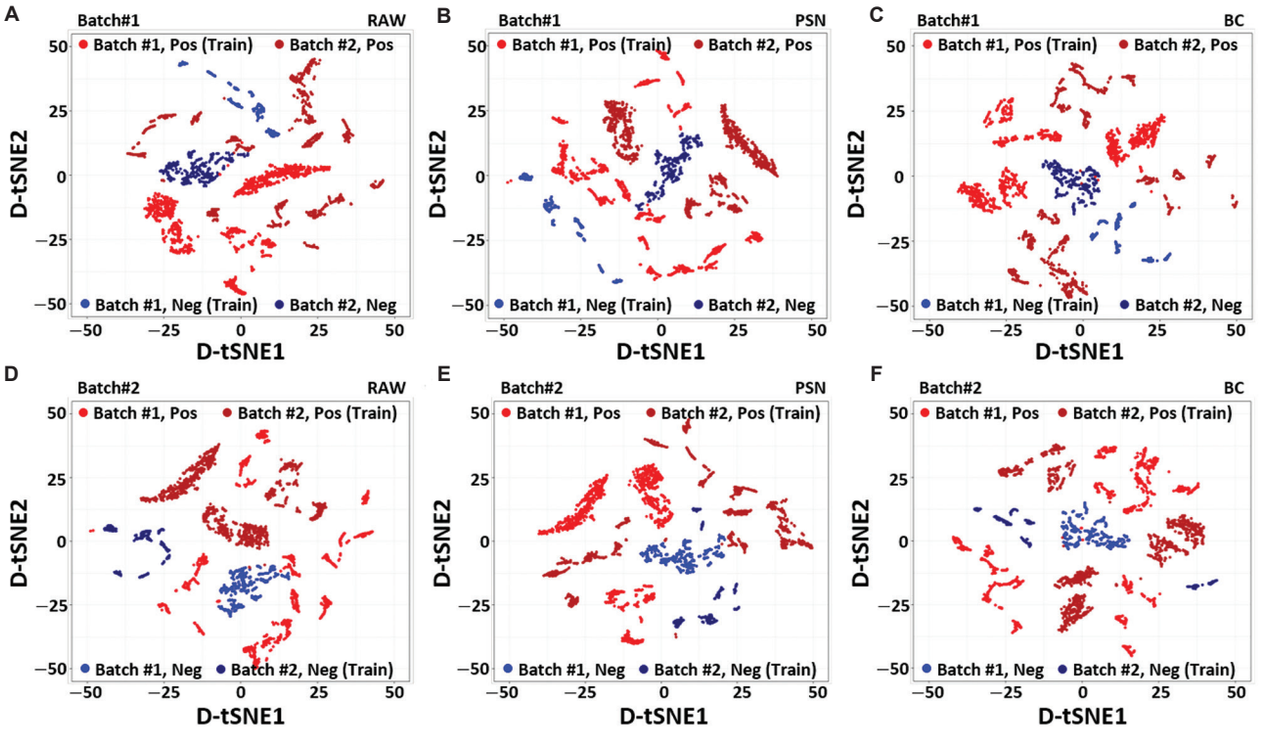

Figure 5. D-tSNE embedding visualization of Pb $(NO_3)_2$ SERS spectra, showing (A, D) RAW, (B, E) PSN, and (C, F) BC preprocessing methods. PCA embeddings were learned using 80% of the training data from one batch, and the remaining 20% were projected using the learned PCA embedding. D-tSNE was then applied for dimensionality reduction, preserving the intrinsic data distribution in the projected feature space. Adapted from Ref[76].

Abbreviations: BC: Baseline correction; D-tSNE: Dynamic t-distributed stochastic neighbor embedding; Pb $(NO_3)_2$: Lead(II) nitrate; PCA: Principal component analysis; PSN: Power spectrum normalization; RAW: Raw data; SERS: Surface-enhanced Raman spectroscopy.

deployed model yielded an R^2 of 0.9933 for Pb^{2+} detection. The model structure and test data are illustrated in Figures 8B and C.

The system was successfully applied for Pb^{2+} detection in three types of environmental and biological samples. First, rock samples from Madan, Croatia, and Panasqueira, Portugal, containing minerals such as sphalerite and galena, were analyzed. The Pb^{2+} content in galena samples reached 37%, consistent with XRF measurements.[99] Second, liver and brain tissues of 16-day-old chicken embryos injected with $Cd(NO_3)_2$ were examined, and the measured Pb^{2+} concentrations differed by <10% from AAS results.[100] Third, 22 artificially spiked plasma samples (Pb^{2+} concentrations ranging from 0 to 6 μg/mL) were tested, with the ANN model achieving an R^2 of 0.993 for Pb^{2+} detection.

(b) Square wave anodic stripping voltammetry-machine learning-based interference-resistant technology

In agricultural soils, the coexistence of Cu^{2+} and Pb^{2+} may generate Cu-Pb alloys that suppress Pb^{0} dissolution, while Zn^{2+} competes for electrode active sites, resulting in square wave anodic stripping voltammetry (SWASV) detection errors exceeding 25%.[75] To address these challenges, Liu *et al*.[75] first employed two-dimensional correlation spectroscopy to explore multi-ion interference mechanisms. The analysis of the synchronous and asynchronous spectra revealed that the interference intensity followed the descending order of Cu^{2+}, Zn^{2+}, and Cd^{2+}, suggesting that variations in the Pb^{2+} peak current occurred earlier than those associated with Zn^{2+}. To overcome the limited information provided by peak currents, they selected 98 characteristic dissolution currents from a total of 281 full-range currents (−1.2 to 0.2 V) through an RF algorithm.[75,102] These currents captured both target ion peaks and background interference.

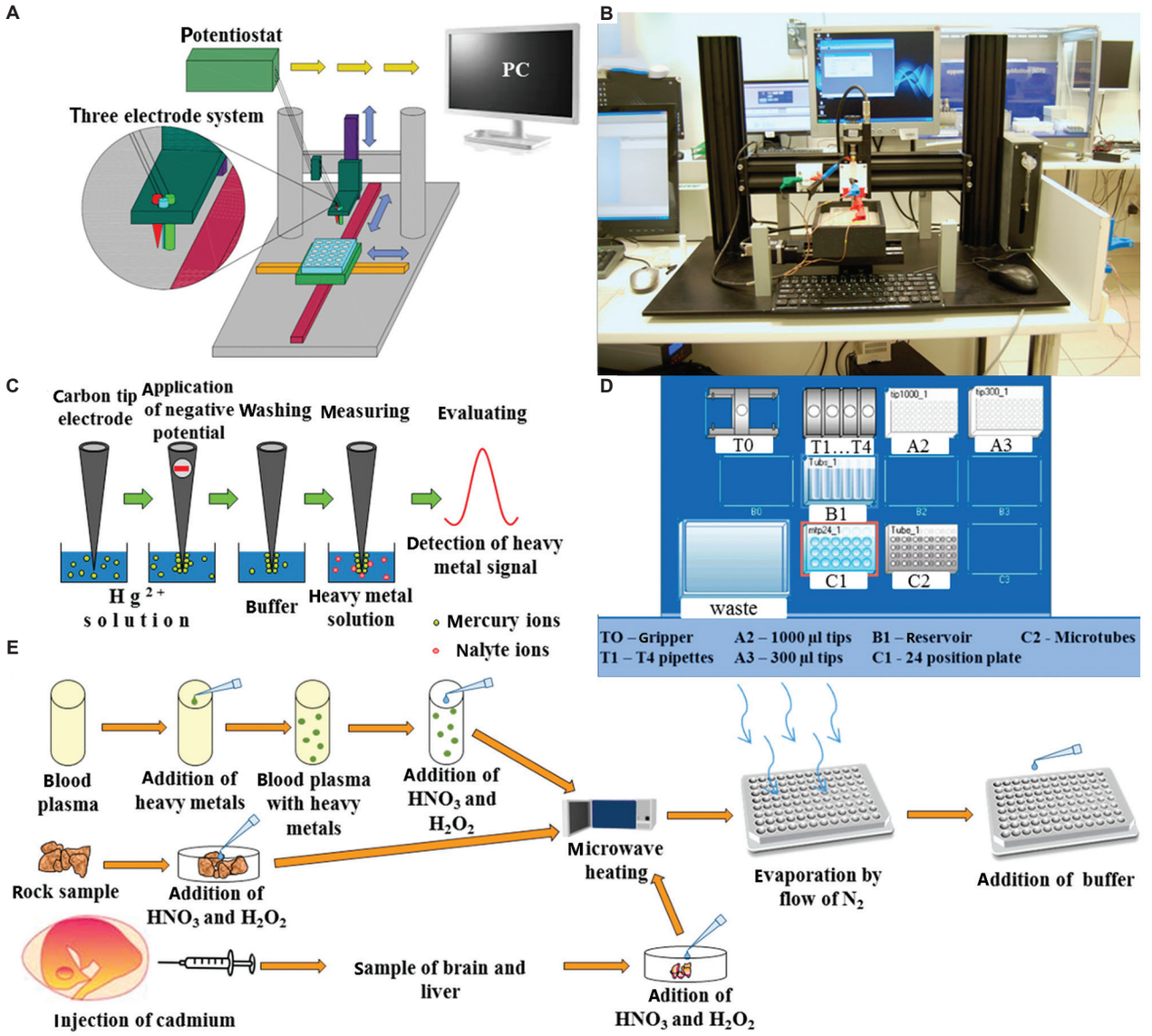

Figure 6. (A) Schematic illustration of the robotic platform equipped with a three-electrode electrochemical detection system. (B) Photograph of the robotic platform. (C) Stepwise preparation of a thin-film mercury layer on a carbon-tip electrode. (D) Schematic of the automated pipetting system (epMotion 5075) used to add buffer solutions to mineralized rock samples. (E) Preparation of plasma, stone, and chicken embryo organ samples before heavy metal detection. Adapted from Ref[77].

Based on these features, feature-RF (for Cd^{2+}) and feature-SVR (for Pb^{2+}) models were developed.

The feature-SVR model for Pb^{2+} detection employed an RBF kernel, with optimal parameters determined via PSO. On the validation set, the feature-SVR model for Pb^{2+} detection achieved a validation set coefficient of determination Rv^2 of 0.992 and an RMSE of the validation set (RMSEV) of 5.348 µg/L. Compared with the conventional peak–SVR model (which had an RMSEV of 25.119 µg/L for Pb^{2+} detection), this represents an approximately 4.70-fold improvement in detection accuracy (calculated based on the reduction in RMSEV). For practical evaluation, agricultural soil extracts from two Chinese provinces (Zn^{2+} 100–250 µg/L, Cu^{2+} 50–125 µg/L) were measured. SWASV signals were collected and input into the feature-SVR model, producing Pb^{2+} results with deviations of <5% relative to ICP-MS and recovery rates of 107.10%. Figure 8 depicts the workflow of feature–SVR-based Pb^{2+} detection applied to soil extracts.

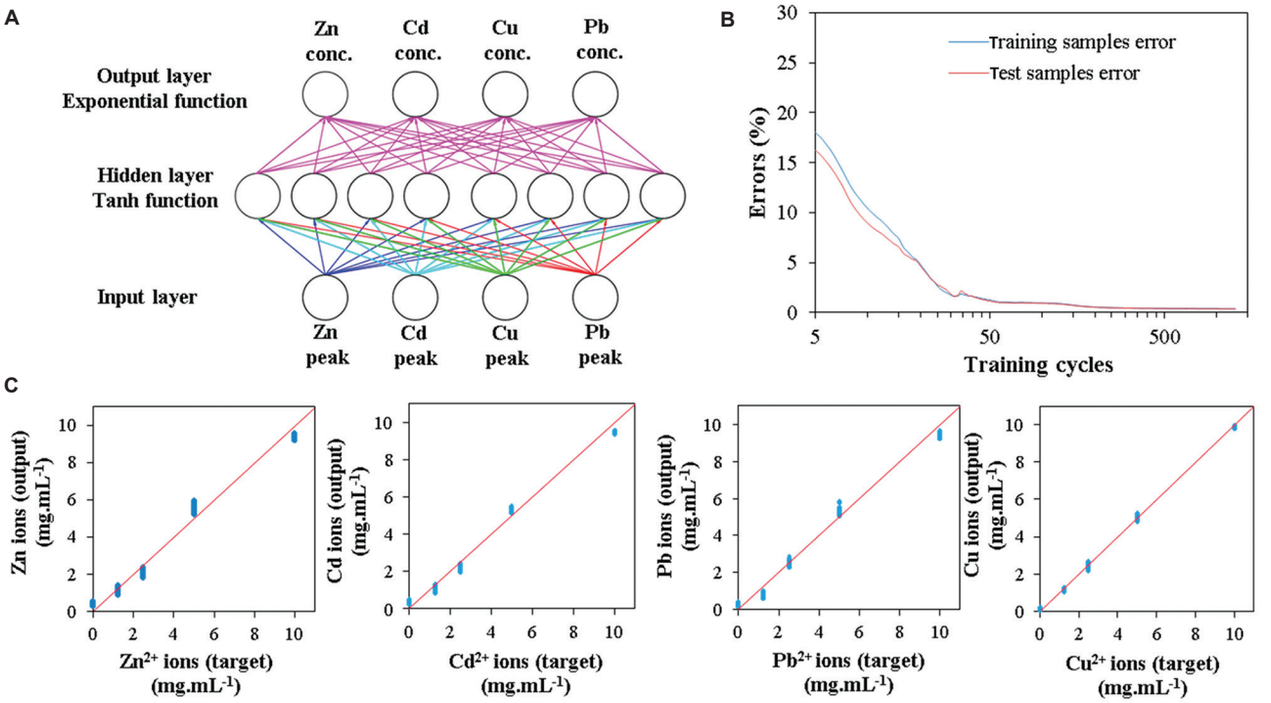

Figure 7. (A) Design of the finalized custom neural network model. (B) Network training incorporating early stopping to prevent overfitting. (C) Goodness-of-fit between test targets and artificial neural network predictions. Adapted from Ref[101].

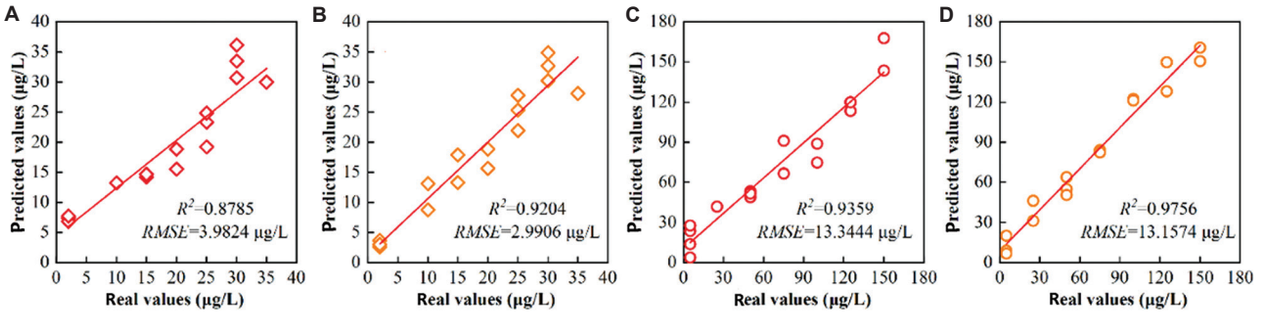

Figure 8. Validation results of the model for (A, B) $Cd^{2+}$and (C, D) Pb^{2+} concentration measurements. Red markers represent model inputs derived from manually extracted absorbance (A_{manu}), whereas orange markers represent inputs obtained from algorithm-extracted absorbance (A_{algo}). Adapted from Ref[89].
Abbreviation: RMSE: Root mean square error.

(c) Automatic peak-area extraction algorithm coupled with support vector regression for square wave anodic stripping voltammetry

Ye *et al.*[89] sought to solve a key limitation of conventional SWASV, where peak height signals are highly susceptible to potential shifts caused by soil pH variations. They proposed replacing peak height with the dissolution peak area, which physically represents the total ion charge and exhibits stronger resistance to potential drift. A custom algorithm facilitated the automatic extraction of peak areas (A_{algo}) for Pb^{2+}, Zn^{2+}, Cd^{2+}, Bi^{3+}, and Cu^{2+}, combining Savitzky–Golay (S–G) smoothing with the peakdet function for extreme point detection and iterative slope-matrix processing.

Using A_{algo} as input, an A_{algo}–SVR model was constructed. For Pb^{2+} detection in soils, the model attained an R^2 of 0.98 and an RMSE of 13.12 μg/L, representing a 34% improvement over the manually extracted peak height-based I_{manu}–SVR model (R^2 = 0.88, RMSE = 200 μg/L). The model demonstrated higher stability, with only a 2.69 μg/L

difference in RMSE between the calibration and validation sets. Model validation results for Cd^{2+} and Pb^{2+} concentrations are illustrated in Figures 8A and B. In applications to various agricultural soils, this approach showed strong adaptability. In southern red soils (pH 4.8–5.2) and northern saline soils (pH 7.2–7.8), the predicted Pb^{2+} concentrations closely matched the measured values, with deviations from ICP-MS results ranging from 0.18 to 0.47 μg/L and an RSD of 1.83%, as shown in Figure 8C and D.

(d) Potentiometric sensor array combined with an ANNs for simultaneous multi-ion detection

Wilson *et al*.[81] engineered an array of nine polyvinyl chloride membrane potentiometric sensors, comprising four Pb^{2+}-selective sensors, two Zn^{2+} sensors, one Cu^{2+} sensor, one Cd^{2+} sensor, and two general-purpose sensors. By leveraging sensor cross-sensitivity and integrating their outputs with an ANN, simultaneous multi-ion detection was achieved. The sensor array exhibited a response range for Pb^{2+} from 1.0×10^{-6} to 1.0×10^{-2} mol/L, with slopes close to the Nernstian response (28.87–30.42 mV/dec). The ANN was trained using Bayesian regularization, and the optimal network topology was determined to be "9–6–4" (input-hidden-output layers). For the external test set, the predicted Pb^{2+} concentrations yielded an R^2 value of 0.99, with a slope of 1.03 and an intercept of -6.39×10^{-3}, consistent with previously reported findings.

In the analysis of real soil samples, this technique was applied to eight roadside soil samples collected from suburban areas of Barcelona, Spain. The samples were digested using a mixture of 35% hydrochloric acid and 65% nitric acid in a 3:1 (v/v) ratio. The resulting solutions were analyzed using the potentiometric sensor array to obtain electromotive force signals, which were then processed by the pre-trained ANN model to estimate Pb^{2+} concentrations. The predicted values were subsequently compared with AAS measurements. The results demonstrated a linear relationship between the Pb^{2+} concentrations determined by the sensor array and the AAS reference values, described by the regression equation $Y = 1.07\ (\pm 0.14)\ X + 2.65 \times 10^{-2}\ (\pm 5.98 \times 10^{-2})$, with $R^2 = 0.950$. Paired *t*-test analysis yielded a calculated *t*-value of 0.12, which is lower than the critical value ($t^* = 2.36$) at seven degrees of freedom and a 95% confidence level. The correlation between Pb^{2+} concentrations in soil samples and the reference values is depicted in Figure 9. The detected Pb^{2+} concentrations ranged from 6.60 to 36.10 mg/kg, with an RSD of ±3.70 mg/kg, well below the maximum permissible limit of 300 mg/kg set by EU Directive 86/278/EEC.

5.1.3. Smartphone-coupled colorimetric and fluorescence single-modality detection

The colorimetric- or fluorescence-based smartphone-coupled single-modal detection technique combines the visual advantages of colorimetric or fluorescent sensors with the portability and accessibility of smartphones. By employing AI algorithms to convert visual signals into Pb^{2+} concentration data, this technique meets the demand for rapid, on-site qualitative and quantitative analysis of soil Pb^{2+}. This method does not require large-scale instrumentation and is well-suited for use in local agricultural settings.

(a) Fluorescence-smartphone dual-modal technology based on polyimide covalent organic framework

Yu *et al*.[74] synthesized a novel polyimide covalent organic framework (COF) that exhibits strong fluorescence at 496 nm on excitation at 436 nm. The presence of Pb^{2+} reduces fluorescence intensity through a combination of fluorescence resonance energy transfer, photoinduced electron transfer, and dynamic quenching mechanisms, while simultaneously inducing a visible color change in the solution from yellow to green.[9] Red–green–blue (RGB) images were acquired using a smartphone at a fixed working distance of 30.00 ± 0.50 cm. The operating principle of the fluorescence-smartphone dual-modal platform is schematically represented in Figure 10.

Yu *et al*.[74] employed a Lasso regression model, utilizing the rate of fluorescence intensity change and the (G + B)/R ratio as input variables, with Pb^{2+} concentrations ranging from 0.10 nM to 1 μM as the output. In the analysis of Dianchi Lake water samples, the model achieved recoveries of 104.00%, 102.10%, and 101.9% for 1, 5, and 10 nM Pb^{2+}, respectively, with an RSD below 6.73%. The LOD was 50 pM, which is well below the United States Environmental Protection Agency drinking water limit of 15 μg/L.

In addition, a swab-type sensor, prepared by coating the COF onto a medical swab, was applied for detecting Pb^{2+} residues on the floor of an electronics factory. Following direct surface swabbing, the collected samples were analyzed via smartphone-based imaging, achieving 90% accuracy at 0.10 μM

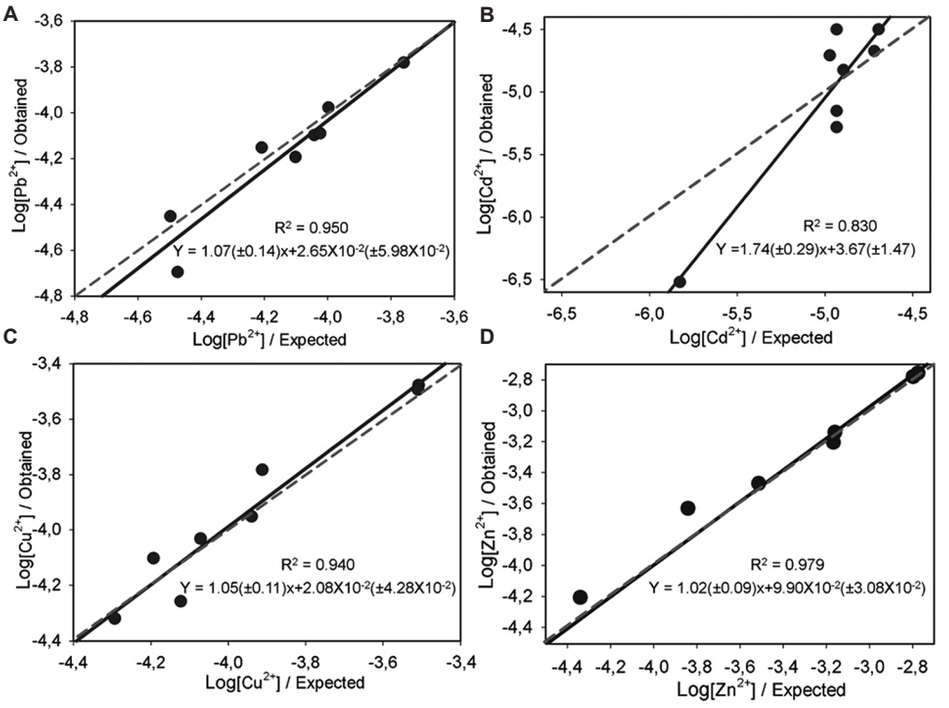

Figure 9. Correlation between measured and reference concentrations of (A) Pb^{2+}, (B) Cd^{2+}, (C) Cu^{2+}, and (D) Zn^{2+} in real soil samples. The dashed lines represent the ideal values, whereas the solid lines represent linear regression fits of the measured data. Uncertainty intervals were calculated at the 95% confidence level. Adapted from Ref.[81]. Copyright © 2012 WILEY-VCH Verlag GmbH and Co. Reprinted with permission of WILEY-VCH Verlag GmbH and Co.

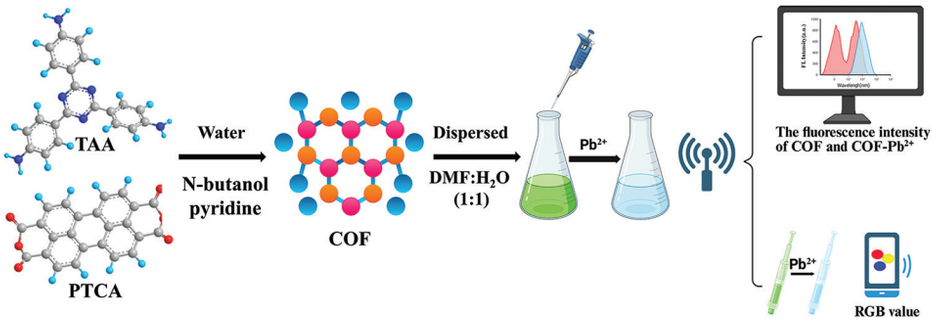

Figure 10. Schematic representation of the fluorescence-smartphone dual-modal platform. Image created by the authors.

Abbreviations: COF: Covalent organic framework; DMF: N,N-dimethylformamide; PTCA; PTCA: Perylene-3,4,9,10-tetracarboxylic acid;; RGB: Red-green-blue; TAA: Thioacetamide.

Pb^{2+}. Compared with conventional portable fluorometers, which require laboratory analysis and approximately 1 h per measurement, this approach reduced analysis time by approximately 80%.

(b) Guanine-functionalized gold nanoparticle-based colorimetric smartphone technology

Sajed *et al.*[73] employed guanine-functionalized gold nanoparticles (GNPs; diameter 18 ± 1.20 nm) as colorimetric probes. The presence of Pb^{2+} induced GNP aggregation, resulting in a distinct color change of the solution from red (absorption peak at 525 nm) to purple (absorption peak at 575 nm). Images acquired using a smartphone (Samsung Galaxy A3) were processed using a custom-developed application to extract RGB features. To address the non-linear relationship between RGB signals and

Pb^{2+} concentration, a non-linear regression model ($\ln[Pb^{2+}\text{ concentration}] = \beta_i X_{ij}$) was constructed using 18 feature variables, achieving an LOD of 0.5 ppb over a concentration range of 0.50–2,000 ppb.

The system achieved recoveries of 86% (RSD = 2.70%) for 0.15 ppb Pb^{2+} spiked in tap water from Tehran, Iran, and 93% (RSD = 2.40%) for 15 ppb Pb^{2+} spiked in Zamzam water from Mecca, Saudi Arabia, with deviations from ICP-MS results below 5%. The predictive performance of the model was validated in Figure 11, with correlation coefficients exceeding 0.99 for both the training and test datasets. Compared with conventional dithizone-based colorimetric methods, which often exhibit RSDs >10% and require toxic chloroform extraction, this system demonstrates improved accuracy while eliminating the risk of secondary contamination.

5.2. Multi-modal AI fusion detection technology

Single-modal techniques are prone to interference from complex soil matrices, including fluorescence quenching caused by high organic matter and electrode passivation induced by high salinity, as well as the synergistic effects of multiple interfering factors, which reduce detection accuracy. Multi-modal fusion technologies integrate two or more sensing signals based on distinct principles. By exploiting the differential responses of these signals to interference and applying deep learning algorithms such as CNN to extract cross-modal features, multi-modal approaches significantly enhance the anti-interference capability and stability of soil Pb^{2+} detection. This strategy represents a promising approach for addressing challenges associated with complex soil matrices.

A typical example is the fluorescence-electrochemical dual-modal sensor coupled with CNN technology. Wang *et al*.[78] designed a dual-modal sensor that employs etched CdTe/CdS quantum dots as the fluorescent sensing material and gold electrodes modified with sea urchin-like iron oxyhydroxide (FeOOH) as the electrochemical sensing material (Figure 12). In the fluorescent component, Cd^{2+} vacancies were created by ethylenediaminetetraacetic acid etching. Filling these vacancies with Cd^{2+} restores fluorescence, whereas Pb^{2+} interferes by competitively binding to thiol ligands on the quantum dot surface. In the electrochemical component, the unique FeOOH structure enhances Pb^{2+} adsorption and produces stripping peaks within the potential range of −0.3 to 0.1 V. The modification of sensing materials and the detection procedure of the dual-modal sensor are illustrated in the one-dimensional CNN model, which was designed to predict Pb^{2+} concentrations (10–49.10 μg/L) using fluorescence spectra (450–680 nm) and electrochemical current curves. The model was trained on 44 sets of simulated water samples containing 10 mg/L humic acid, producing an R^2 of 0.999 and a mean absolute error (MAE) of 0.60 μg/L. This performance represents an approximately 13-fold improvement over a linear regression model (MAE = 8.22 μg/L).

In soil leachates containing 10 mg/L humic acid (Pb^{2+} 25 μg/L), the CNN model enabled direct analysis without pre-treatment, providing a recovery

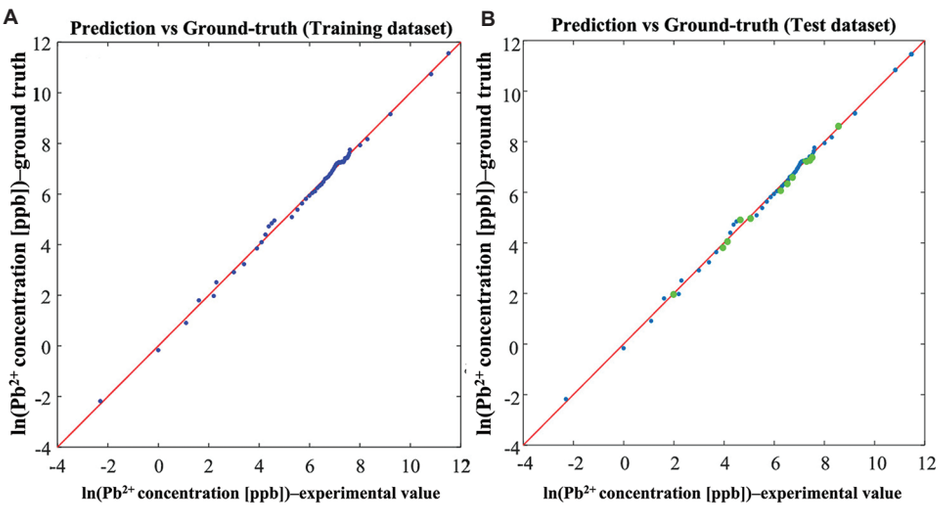

Figure 11. Prediction versus ground-truth of (A) training and (B) test datasets for Pb^{2+} ppb. Adapted from Ref.[73]. Copyright © 2020 American Chemical Society. Reprinted with permission from American Chemical Society

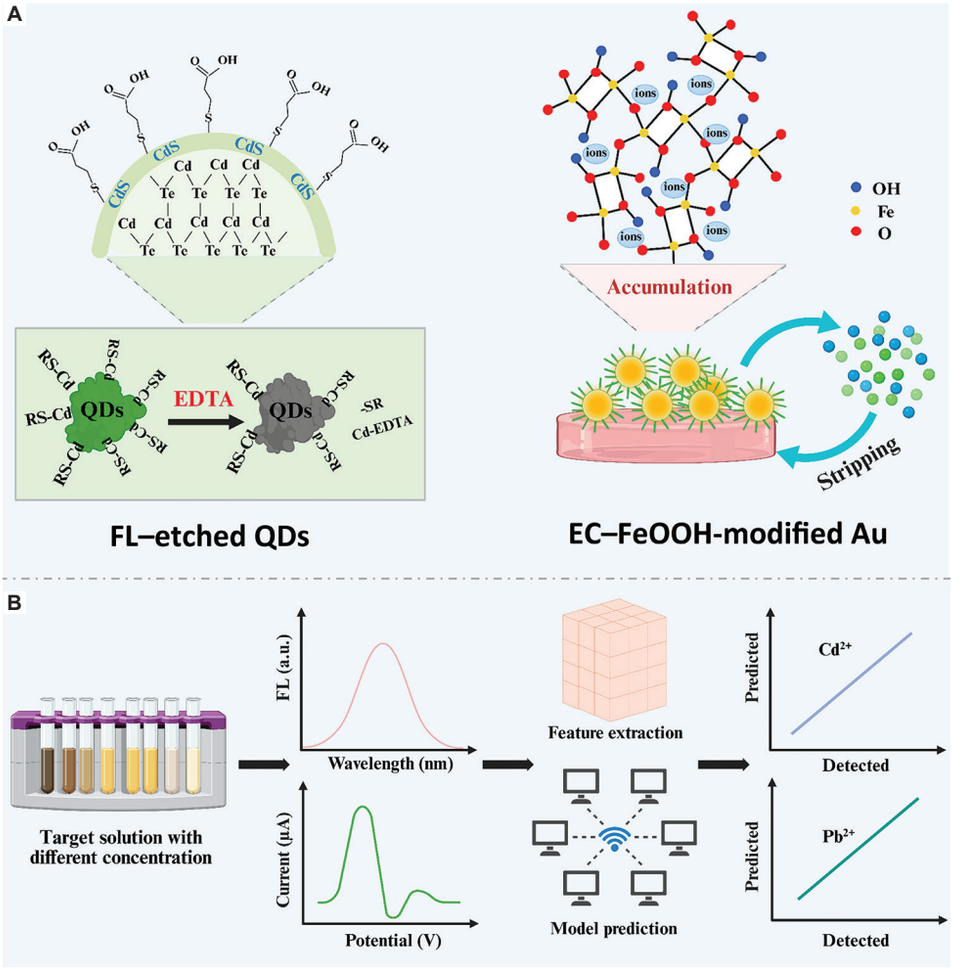

Figure 12. (A) Modification of sensing materials for the FL–EC dual-modal sensor. (B) Schematic representation of the analytical procedure for heavy metal ion detection. Image created by the authors.
Abbreviations: EC: Electrochemical; EDTA: Ethylenediaminetetraacetic acid; FeOOH: Iron oxyhydroxide; FL–EC: Fluorescence-electrochemical; QD: Quantum dot.

of 98.60% with an RSD of 4.27%. By comparison, single fluorescent sensors were adversely affected by fluorescence quenching and competitive interference from Pb^{2+}, commonly exhibiting recoveries of only 75–85%.

For seawater samples (salinity 35%, Pb^{2+} 30 μg/L), the CNN model employed selective feature filtering to mitigate the influence of chloride ions on the electrochemical signal. The predicted results deviated by <4% from ICP-MS measurements, whereas conventional electrochemical sensors typically showed REs exceeding 20% due to electrode passivation induced by high salinity.

5.3. Automated integrated AI-based detection system

Most of the single- and multi-modal techniques described above still rely heavily on laboratory operations, such as manual sample pre-treatment and instrument calibration, which limits their suitability for rapid, multi-site analysis of agricultural soils. Automated integrated systems address these limitations by combining modules for sample preparation, detection, and data analysis, while leveraging AI algorithms to enable fully automated workflows and real-time data interpretation. This approach substantially reduces analysis time and minimizes human intervention, serving as a key platform for translating AI-based Pb^{2+} detection from the laboratory to in-field applications.

An illustrative example is a smartphone-controlled automated integrated detection system. Liu *et al.*[88] developed a smartphone-operated device that integrates automated sample pre-treatment, SWASV detection, and machine learning analysis. This system enables rapid,

in-field determination of Pb^{2+} in agricultural soils. The core workflow is schematically illustrated in Figure 13, showing the entire process from soil sampling to Pb^{2+} concentration output without manual intervention.

The procedure includes ultrasound-assisted extraction using an acetate buffer, vacuum filtration, and ultraviolet photolysis of dissolved organic matter. Ultrasound-assisted extraction was performed at 150 W for 30 min using a 0.30 M acetate buffer (pH 5.0). This approach enhanced extraction efficiency by 32-fold relative to the conventional 16 h BCR method. The weakly acidic soluble Pb^{2+} fraction reached 17.13%, significantly higher than the 11.85% obtained with the BCR procedure.

The device and its smartphone application interface demonstrate an integrated automated system for the detection of heavy metals in soil. The hardware incorporates an ultrasonic processor, vacuum pump, ultraviolet photolysis module, and a custom-built potentiostat. The application enables remote operation of both sample pre-treatment and detection.

The AI component of this technology is centered on the combination of a peak information acquisition algorithm and machine learning models. The algorithm first removes noise using Savitzky-Golay filtering and then uses the Bi^{3+} peak as a reference to correct for peak position shifts caused by potential drift. It automatically identifies the peak heights and widths of Zn^{2+}, Cd^{2+}, Pb^{2+}, and Cu^{2+}. Even when Cu^{2+} concentrations are below 25 μg/L and shoulder peaks are formed, the algorithm accurately recognizes them using first-derivative feature points. The extracted peak parameters demonstrate strong consistency with manually obtained values, with $R^2 = 0.99$ for peak heights and $R^2 = 0.94$ for peak widths. Based on these parameters, an H-W-SVR model was constructed using peak heights and widths as input variables.

For Pb^{2+} in agricultural soils, the model achieved a validation set R^2 of 0.94 and an RMSEV of 17.73 μg/L. This represents a 5.40% improvement in stability relative to the H-SVR model using only peak heights as input (R^2 = 0.89, RMSEV = 23.27 μg/L). In practical applications, the device was tested on simulated contaminated samples of eight representative Chinese soils, including red soil, black soil, and tidal soil, each spiked with 200 μg/L Pb^{2+}. The measured Pb^{2+} concentrations showed no significant difference from those obtained using the BCR-ICP-MS method (t = 0.040 < critical value 1.76), with spiked recoveries reaching 105.19 ± 11.47%. Individual sample analysis required only 75 min, corresponding to a 13.60-fold improvement in efficiency compared with the BCR-ICP-MS procedure, which takes 17 h per sample.

In summary, AI has been integrated with optical, electrochemical, and colorimetric/fluorescent sensors to detect Pb^{2+} in agricultural soils. The framework includes single-modality optimization, multi-modal fusion, and fully automated processes. It enhances sensitivity, anti-interference capability, and efficiency through precise signal analysis. Some methods adapted to soil leachates have shown preliminary field applicability, supporting farmland contamination screening.

However, several limitations remain. One major issue is limited adaptation to diverse soil matrices, as most systems originate from aqueous Pb^{2+} detection and do not fully address interferences such as clay adsorption or organic matter complexation. Another

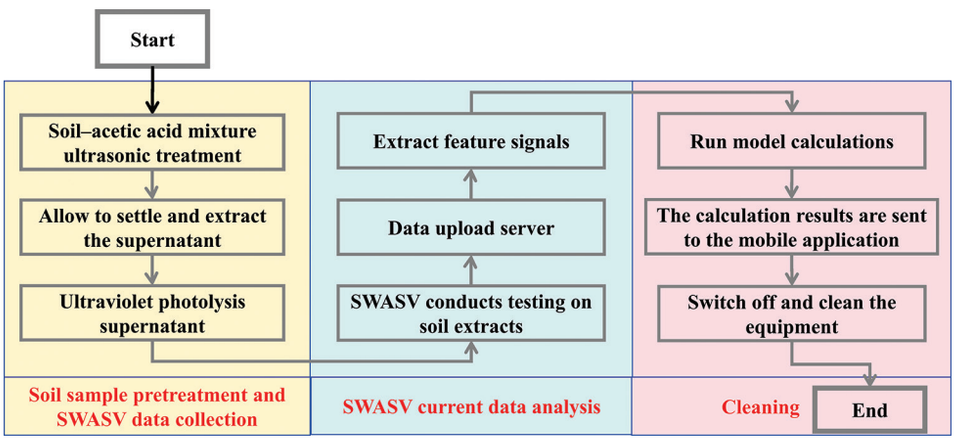

Figure 13. Workflow for the determination of weakly acid–soluble heavy metals in soil using the newly designed integrated automated device. Image created by the authors.
Abbreviation: SWASV: Square wave anodic stripping voltammetry.

concern is the lack of *in situ* detection, as many systems rely on ultrasonic extraction or filtration. A further challenge is the limited interpretability of AI models, which provides little insight into soil–Pb interactions or interference signals, thereby restricting deployment in complex farmland environments.

6. Conclusion and future directions

6.1. Conclusion

This review systematically examines AI-based techniques for Pb^{2+} detection in agricultural soils. Traditional "sample pre-treatment-instrumental analysis" methods are accurate but labor-intensive, require laboratory settings, and are prone to matrix interference or Pb volatilization. In contrast, AI-based techniques efficiently extract information from spectroscopic and electrochemical signals, overcoming limitations. Their applications can be categorized into three main strategies:

(i) Single-modality models, which improve sensitivity and specificity.
(ii) Multi-modal fusion methods, which mitigate interference from complex soil matrices.
(iii) Fully automated platforms, which enable rapid, field-deployable detection.

Together, these strategies form an integrated framework from signal optimization to practical deployment, supporting efficient farmland Pb^{2+} screening. Despite progress, challenges remain. Existing AI detection prototypes are largely based on aqueous Pb^{2+} studies and, when applied to soils, primarily focus on leachates, insufficiently addressing matrix-specific interferences. In addition, optimization for diverse soils is limited.

Automated systems reduce pre-treatment effort but still rely on ultrasonic extraction or vacuum filtration, preventing true *in situ* measurements and limiting the monitoring of spatial heterogeneity. Furthermore, AI models mainly predict Pb^{2+} concentrations with limited interpretability of soil-Pb interactions or interference signals, restricting their applicability in complex farmland conditions.

6.2. Future directions

Future research on Pb^{2+} detection should emphasize synergistic strategies integrating materials, AI, and field applications. At the material level, amino-functionalized COF probes targeting exchangeable Pb^{2+} enable selective detection, while CNN-based automatic feature extraction enhances weak signals. Degradable nanosensor arrays, coupled with unmanned aerial vehicle remote sensing and ground-based AI networks, can facilitate multi-scale mapping from field plots to rhizosphere microzones. At the AI level, hybrid models that combine physical constraints with data-driven learning can embed stripping kinetics and SERS characteristics, and leverage two-dimensional correlation spectroscopy interference sequences to quantify feature contributions, thereby advancing both predictive accuracy and interpretability. For field applications, assembling specialized training sets across diverse soil types and chemical forms, alongside dynamic updates via Internet-of-Things devices and federated learning, can strengthen model stability and generalization under complex environmental conditions.

Acknowledgments

None.

Funding

This research was supported by the National Natural Science Foundation of China (No. 5180425), the Special Project for Young Scientific and Technological Talents of Nanchong Science and Technology Bureau (25KJCXRC0007), and the Science and Technology Bureau of Xiangtan City, Hunan Province (CG-YB20240007).

Conflict of interest

The authors declare that they have no competing interests.

Author contributions

Conceptualization: Zhenxing Wang, Kefan Yang, Hui Xiao
Formal analysis: Zhenxing Wang
Investigation: Haolei Gong
Methodology: Zhenxing Wang, Kefan Yang, Wei Zou, Hui Xiao
Writing–original draft: Zhenxing Wang
Writing–review & editing: Haolei Gong, Wei Zou, Hui Xiao

Availability of data

Not applicable.

References

1. Banwart SA, Nikolaidis NP, Zhu YG, Peacock CL, Sparks DL. Soil functions: Connecting earth's critical zone. *Annu Rev Earth Planet Sci*. 2019;47(1):333-359. doi: 10.1146/annurev-earth-063016-020544
2. Kafle G. Abiotic and biotic factors affecting formation of soil in the Earth. *J Agric Nat Resour*. 2023;6(1):20-31. doi: 10.3126/janr.v6i1.71850
3. Spiridonov V, Ćurić M, Novkovski N. Biosphere: Ecosystem diversity and environmental change. In: *Atmospheric Perspectives: Unveiling Earth's Environmental Challenges*. Berlin: Springer; 2025. p. 51-81. doi: 10.1007/978-3-031-86757-6_4
4. Padhiary M, Kumar R. Assessing the environmental impacts of agriculture, industrial operations, and mining on agro-ecosystems. In: *Smart Internet of Things for Environment and Healthcare*. Berlin: Springer; 2024. p. 107-126. doi: 10.1007/978-3-031-70102-3_8
5. Sodango TH, Li X, Sha J, Bao Z. Review of the spatial distribution, source and extent of heavy metal pollution of soil in China: Impacts and mitigation approaches. *J Health Pollut*. 2018;8(17):53. doi: 10.5696/2156-9614-8.17.53
6. Rakshith BL, Gautam S, Asirvatham LG, *et al*. Biofilm-associated microplastic contamination in rural soil and water: Emerging hazards to ecosystems. *Sci Total Environ*. 2025;1004:180806. doi: 10.1016/j.scitotenv.2025.180806
7. Rashmi I, Roy T, Kartika K, *et al*. Organic and inorganic fertilizer contaminants in agriculture: Impact on soil and water resources. In: *Contaminants in Agriculture: Sources, Impacts and Management*. Berlin: Springer; 2020. p. 3-41. doi: 10.1007/978-3-030-41552-5_1
8. Lu D, Zhang C, Zhou Z, *et al*. Pollution characteristics and source identification of farmland soils in Pb-Zn mining areas through an integrated approach. *Environ Geochem Health*. 2023;45(5):2533-2547. doi: 10.1007/s10653-022-01355-0
9. Rathikannu S, Gautam S, Joshi SK, *et al.* FTIR based assessment of microplastic contamination in soil water and insect ecosystems reveals environmental and ecological risks. *Sci Rep*. 2025;15(1):28615. doi: 10.1038/s41598-025-14507-w
10. Huang Y, Wang L, Wang W, Li T, He Z, Yang X. Current status of agricultural soil pollution by heavy metals in China: A meta-analysis. *Sci Total Environ*. 2019;651:3034-3042. doi: 10.1016/j.scitotenv.2018.10.185
11. Jomova K, Alomar SY, Nepovimova E, Kuca K, Valko M. Heavy metals: Toxicity and human health effects. *Arch Toxicol*. 2025;99(1):153-209. doi: 10.1007/s00204-024-03903-2
12. Xie Y, Zhou G, Huang X, *et al*. Study on the physicochemical properties changes of field aging biochar and its effects on the immobilization mechanism for Cd2+ and Pb2+. *Ecotoxicol Environ Saf*. 2022;230:113107. doi: 10.1016/j.ecoenv.2021.113107
13. Huihui Z, Xin L, Zisong X, *et al*. Toxic effects of heavy metals Pb and Cd on mulberry (*Morus alba* L.) seedling leaves: Photosynthetic function and reactive oxygen species (ROS) metabolism responses. *Ecotoxicol Environ Saf*. 2020;195:110469. doi: 10.1016/j.ecoenv.2020.110469
14. Collin MS, Venkatraman SK, Vijayakumar N, *et al.* Bioaccumulation of lead (Pb) and its effects on human: A review. *J Hazard Mater Adv*. 2022;7:100094. doi: 10.1016/j.hazadv.2022.100094
15. Kumar A, Kumar A, Cabral-Pinto MMS, *et al*. Lead toxicity: Health hazards, influence on food chain, and sustainable remediation approaches. *Int J Environ Res Public Health*. 2020;17(7):2179. doi: 10.3390/ijerph17072179
16. Meena V, Dotaniya ML, Saha JK, Das H, Patra AK. Impact of lead contamination on agroecosystem and human health. In: *Lead in Plants and the Environment*. Berlin: Springer; 2019. p. 67-82. doi: 10.1007/978-3-030-21638-2_4
17. Tawfik W, El-Saeed M, Khalil A, Fikry M. Detection of heavy metal elements by using advanced optical techniques. *J Egypt Soc Basic Sci Phys*. 2024;1(1):99-127. doi: 10.21608/jesbsp.2024.255191.1002
18. Yan P, Lin K, Wang Y, *et al*. Spatial interpolation of red bed soil moisture in Nanxiong basin, South China. *J Contam Hydrol*. 2021;242:103860. doi: 10.1016/j.jconhyd.2021.103860
19. Murugasridevi K, Mageshen V, Poornima R, Ramya A. The role of robotics in sustainable agriculture and waste management. In: *Addressing Environmental Challenges with AI, Robotics, and Augmented Reality*. United States: IGI Global Scientific Publishing; 2025. p. 47-82. doi: 10.4018/979-8-3373-1892-9.ch003
20. Baydarashvili M, Shrednik N, Spasovskaia A. Detection method of pollution with heavy metals ions of the soil. *Procedia Eng*. 2017;189:630-636. doi: 10.1016/j.proeng.2017.05.100
21. Chinnaiyan B, Balasubaramanian S, Jeyabalu M, Warrier GS. AI applications-computer vision and natural language processing. In: *Model Optimization Methods for Efficient and Edge AI: Federated Learning Architectures, Frameworks and Applications*. New Jersey: Wiley; 2025. p. 25-41. doi: 10.1002/9781394219230.ch2
22. Ranjith J. Enhancing stock market trend prediction using explainable artificial intelligence and multi-source data. *Fusion Pract Appl*. 2024;16(2):178-189. doi: 10.54216/FPA.160211
23. Sharafat MS, Kabya ND, Emu RI, *et al*. An IoT-enabled

AI system for real-time crop prediction using soil and weather data in precision agriculture. *Smart Agric Technol*. 2025;12:101263.
doi: 10.1016/j.atech.2025.101263

24. Misra NN, Dixit Y, Al-Mallahi A, Bhullar MS, Upadhyay R, Martynenko A. IoT, big data, and artificial intelligence in agriculture and food industry. *IEEE Internet Things J*. 2020;9(9):6305-6324.
doi: 10.1109/JIOT.2020.2998584
25. Tavares TR, Minasny B, McBratney A, *et al*. Estimating plant-available nutrients with XRF sensors: Towards a versatile analysis tool for soil condition assessment. *Geoderma*. 2023;439:116701.
doi: 10.1016/j.geoderma.2023.116701
26. Tan K, Ma W, Wu F, Du Q. Random forest-based estimation of heavy metal concentration in agricultural soils with hyperspectral sensor data. *Environ Monit Assess*. 2019;191(7):446.
doi: 10.1007/s10661-019-7510-4
27. Derakhshan Nejad Z, Jung MC, Kim KH. Remediation of soils contaminated with heavy metals with an emphasis on immobilization technology. *Environ Geochem Health*. 2018;40(3):927-953.
doi: 10.1007/s10653-017-9964-z
28. Liu G, Liao B, Lu T, *et al*. Insight into immobilization of Pb2+ in aqueous solution and contaminated soil using hydroxyapatite/attapulgite composite. *Colloids Surf A Physicochem Eng Asp*. 2020;603:125290.
doi: 10.1016/j.colsurfa.2020.125290
29. Bian H, Ma Y, Ji P, *et al*. Effect of enzyme-induced carbonate precipitation (EICP) combined with biochar on lead-contaminated soil solidification and plant growth. *J Environ Chem Eng*. 2025;13:116977.
doi: 10.1016/j.jece.2025.116977
30. Huang S, Yamaji N, Ma JF. Metal transport systems in plants. *Annu Rev Plant Biol*. 2024;75:1-25.
doi: 10.1146/annurev-arplant-062923-021424
31. Rashid A, Schutte BJ, Ulery A, *et al*. Heavy metal contamination in agricultural soil: Environmental pollutants affecting crop health. *Agronomy*. 2023;13(6):1521.
doi: 10.3390/agronomy13061521
32. Hou D, O'Connor D, Igalavithana AD, *et al*. Metal contamination and bioremediation of agricultural soils for food safety and sustainability. *Nat Rev Earth Environ*. 2020;1(7):366-381.
doi: 10.1038/s43017-020-0061-y
33. Silva KN, Paiva SS, Souza FL, Silva D, Martínez-Huitle CA, Santos EV. Applicability of electrochemical technologies for removing and monitoring Pb2+ from soil and water. *J Electroanal Chem*. 2018;816:171-178.
doi: 10.1016/j.jelechem.2018.03.051
34. Fayiga AO, Nwoke O. Phosphate rock: Origin, importance, environmental impacts, and future roles. *Environ Rev*. 2016;24(4):403-415.
doi: 10.1139/er-2016-0003
35. Yang Q, Li Z, Lu X, Duan Q, Huang L, Bi J. A review of soil heavy metal pollution from industrial and agricultural regions in China: Pollution and risk assessment. *Sci Total Environ*. 2018;642:690-700.
doi: 10.1016/j.scitotenv.2018.06.068
36. Zwolak A, Sarzyńska M, Szpyrka E, Stawarczyk K. Sources of soil pollution by heavy metals and their accumulation in vegetables: A review. *Water Air Soil Pollut*. 2019;230(7):164.
doi: 10.1007/s11270-019-4221-y
37. Xia F, Zhao Z, Niu X, Wang Z. Integrated pollution analysis, pollution area identification and source apportionment of heavy metal contamination in agricultural soil. *J Hazard Mater*. 2024;465:133215.
doi: 10.1016/j.jhazmat.2023.133215
38. Tóth G, Hermann T, Da Silva MR, Montanarella L. Heavy metals in agricultural soils of the European Union with implications for food safety. *Environ Int*. 2016;88:299-309.
doi: 10.1016/j.envint.2015.12.017
39. Hwang HM, Fiala MJ, Park D, Wade TL. Review of pollutants in urban road dust and stormwater runoff: Part 1. Heavy metals released from vehicles. *Int J Urban Sci*. 2016;20(3):334-360.
doi: 10.1080/12265934.2016.1193041
40. Andrunik M, Wołowiec M, Wojnarski D, Zelek-Pogudz S, Bajda T. Transformation of Pb, Cd, and Zn minerals using phosphates. *Minerals*. 2020;10(4):342.
doi: 10.3390/min10040342
41. Szwalec A, Mundała P, Kędzior R, Pawlik J. Monitoring and assessment of cadmium, lead, zinc and copper concentrations in arable roadside soils in terms of different traffic conditions. *Environ Monit Assess*. 2020;192(3):155.
doi: 10.1007/s10661-020-8120-x
42. Yan P, Li G, Wang W, Zhao Y, Wang J, Wen Z. Qualitative and quantitative detection of microplastics in soil based on LIF technology combined with OOA-ELM/SPA-PLS. *Microchem J*. 2024;201:110632.
doi: 10.1016/j.microc.2024.110632
43. Sandroni V, Smith C. Microwave digestion of sludge, soil and sediment samples for metal analysis by inductively coupled plasma-atomic emission spectrometry. *Anal Chim Acta*. 2002;468:335-344.
doi: 10.1016/S0003-2670(02)00655-4
44. Jurić M, Golub N, Galić E, Radić K, Maslov Bandić L, Vitali Čepo D. Microwave-assisted extraction of bioactive compounds from mandarin peel: A comprehensive biorefinery strategy. *Antioxidants (Basel)*. 2025;14(6):722.
doi: 10.3390/antiox14060722
45. Pan F, Yu Y, Yu L, *et al*. Quantitative assessment on soil concentration of heavy metal-contaminated soil with various sample pretreatment techniques and detection

methods. *Environ Monit Assess*. 2020;192(12):800. doi: 10.1007/s10661-020-08775-4
46. Perring L, Alonso MI, Andrey D, Bourqui B, Zbinden P. An evaluation of analytical techniques for determination of lead, cadmium, chromium, and mercury in food-packaging materials. *Fresenius J Anal Chem*. 2001;370(1):76-81. doi: 10.1007/s002160100716
47. Patriarca M, Barlow N, Cross A, Hill S, Milde D, Tyson J. Atomic spectrometry update: Review of advances in the analysis of clinical and biological materials, foods and beverages. *J Anal Atomic Spectrom*. 2025;40:541-664. doi: 10.1039/D5JA90008E
48. Li W, Niu N, Guo N, Zhou H, Bu J, Ding A. *Comparative Study on the Determination of Heavy Metals in Soil by XRF and ICP-MS*. England: IOP Publishing; 2021. p. 012075. doi: 10.1088/1742-6596/2009/1/012075
49. Mengtian H, Li X. Quantitative detection of toxic elements in food samples by inductively coupled plasma mass spectrometry (ICP-MS). *Processes*. 2025;13(10):3361. doi: 10.3390/pr13103361
50. Adesina KE, Burgos CJ, Grier TR, Sayam AS, Specht AJ. Ways to measure metals: From ICP-MS to XRF. *Curr Environ Health Rep*. 2025;12(1):7. doi: 10.1007/s40572-025-00473-y
51. Katakam LN, Aboul-Enein HY. Elemental impurities determination by ICP-AES/ICP-MS: A review of theory, interpretation of concentration limits, analytical method development challenges and validation criterion for pharmaceutical dosage forms. *Curr Pharm Anal*. 2020;16(4):392-403. doi: 10.2174/1573412915666190225160512
52. Jin M, Yuan H, Liu B, Peng J, Xu L, Yang D. Review of the distribution and detection methods of heavy metals in the environment. *Anal Methods*. 2020;12(48):5747-5766. doi: 10.1039/D0AY01577F
53. Li Q, Zhang W, Guo Y, Chen H, Ding Q, Zhang L. Oxygenated carbon nanotubes cages coated solid-phase microextraction fiber for selective extraction of migrated aromatic amines from food contact materials. *J Chromatogr A*. 2021;1646:462031. doi: 10.1016/j.chroma.2021.462031
54. Malik LA, Bashir A, Qureashi A, Pandith AH. Detection and removal of heavy metal ions: A review. *Environ Chem Lett*. 2019;17(4):1495-1521. doi: 10.1007/s10311-019-00891-z
55. Wang L, Lamb D, Dong Z, Sanderson P, Du J, Naidu R. Integrating portable X-ray fluorescence site survey and ArcGIS models for rapid risk assessment and mitigation strategies at an abandoned arsenic mine site: A case study. *Environ Technol*. 2025;46(2):266-278. doi: 10.1080/09593330.2024.2354121
56. Hassan MH, Khan R, Andreescu S. Advances in electrochemical detection methods for measuring contaminants of emerging concerns. *Electrochem Sci Adv*. 2022;2(6):e2100184. doi: 10.1002/elsa.202100184
57. Abbruzzini TF, Silva CA, Andrade DA, Carneiro WJO. Influence of digestion methods on the recovery of iron, zinc, nickel, chromium, cadmium and lead contents in 11 organic residues. *Rev Bras Ciên Solo*. 2014;38:166-176. doi: 10.1590/S0100-06832014000100016
58. Mohammed E, Mohammed T, Mohammed A. Optimization of an acid digestion procedure for the determination of Hg, As, Sb, Pb and Cd in fish muscle tissue. *MethodsX*. 2017;4:513-523. doi: 10.1016/j.mex.2017.11.006
59. Wang J, Wang X, Li G, *et al.* Speciation analysis method of heavy metals in organic fertilizers: A review. *Sustainability*. 2022;14(24):16789. doi: 10.3390/su142416789
60. Ratoiu LC, Serafima S, Cortea IM, Duliu OG. A multi-analytical study of a 17th-century wallachian icon depicting the "mother of god with child". *Heritage*. 2023;6(10):6931-6948. doi: 10.3390/heritage6100362
61. Molognoni L, Zarpelon J, de Sá Ploêncio LA, Santos JN, Daguer H. Different approaches for digestion, performance assessment and measurement uncertainty for the analysis of cadmium and lead in feeds. *Food Anal Methods*. 2017;10(6):1787-1799. doi: 10.1007/s12161-016-0718-9
62. Jahanban L, Ebrahimi E, Moradi S, Fallah M, Arani LG, Mohajer R. Estimation of soil specific surface area using some mechanical properties of soil by artificial neural networks. *Environ Monit Assess*. 2018;190(10):614. doi: 10.1007/s10661-018-6980-0
63. Shahare Y, Singh MP, Singh P, *et al.* A comprehensive analysis of machine learning-based assessment and prediction of soil enzyme activity. *Agriculture*. 2023;13(7):1323. doi: 10.3390/agriculture13071323
64. Yang J, Wang X, Wang R, Wang H. Combination of convolutional neural networks and recurrent neural networks for predicting soil properties using Vis-NIR spectroscopy. *Geoderma*. 2020;380:114616. doi: 10.1016/j.geoderma.2020.114616
65. Sharma K, Shivandu SK. Integrating artificial intelligence and Internet of Things (IoT) for enhanced crop monitoring and management in precision agriculture. *Sens Int*. 2024;5:100292. doi: 10.1016/j.sintl.2024.100292
66. Harani P, Gautam S, Joshi SK, Asirvatham LG, Rakshith BL. Advancements in soil moisture estimation through integration of remote sensing and artificial intelligence techniques. *Sci Total Environ*. 2025;1001:180503. doi: 10.1016/j.scitotenv.2025.180503
67. Pyo J, Hong SM, Kwon YS, Kim MS, Cho KH.

Estimation of heavy metals using deep neural network with visible and infrared spectroscopy of soil. *Sci Total Environ*. 2020;741:140162.
doi: 10.1016/j.scitotenv.2020.140162

68. Shi Y, Zhang S, Zhou H, *et al*. Recent developments in heavy metals detection: Modified electrodes, pretreatment methods, prediction models and algorithms. *Metals*. 2025;15(1):80.
doi: 10.3390/met15010080
69. Proshad R, Chandra K, Islam M, *et al*. Evaluation of machine learning models for accurate prediction of heavy metals in coal mining region soils in Bangladesh. *Environ Geochem Health*. 2025;47(5):181.
doi: 10.1007/s10653-025-02489-7
70. Barkhordari MS, Zhou N, Li K, Qi C. Interpretable machine learning for predicting heavy metal removal efficiency in electrokinetic soil remediation. *J Environ Chem Eng*. 2024;12(6):114330.
doi: 10.1016/j.jece.2024.114330
71. Hu T, Chen Q, Lin Z, Qi C, Chai L. Machine learning enables low-cost determination of soil heavy metal concentrations. *ACS ES&T Eng*. 2025;5:3085-3095.
doi: 10.1021/acsestengg.5c00463
72. Mishra A, Kushwaha A, Verma R. Machine learning enabled colorimetric paper strip sensor for the detection of ultra-low concentrations of heavy metal ions. *Sens Actuat A Phys*. 2025:117380.
doi: 10.1016/j.sna.2025.117380
73. Sajed S, Kolahdouz M, Sadeghi MA, Razavi SF. High-performance estimation of lead ion concentration using smartphone-based colorimetric analysis and a machine learning approach. *ACS Omega*. 2020;5(42):27675-27684.
doi: 10.1021/acsomega.0c04255
74. Yu L, Bai L, Liu J, *et al*. Machine learning-assisted fluorescence/smartphone dual-mode platform for lead ion detection using the novel polyimide covalent organic framework. *Talanta*. 2025;295:128346.
doi: 10.1016/j.talanta.2025.128346
75. Liu N, Ye W, Liu G, Zhao G. Improving the accuracy of stripping voltammetry detection of Cd^{2+} and Pb^{2+} in the presence of Cu^{2+} and Zn^{2+} by machine learning: Understanding and inhibiting the interactive interference among multiple heavy metals. *Anal Chim Acta*. 2022;1213:339956.
doi: 10.1016/j.aca.2022.339956
76. Park S, Lee J, Khan S, Wahab A, Kim M. Machine learning-based heavy metal ion detection using surface-enhanced raman spectroscopy. *Sensors (Basel)*. 2022;22(2):596.
doi: 10.3390/s22020596
77. Kudr J, Nguyen HV, Gumulec J, *et al.* Simultaneous automatic electrochemical detection of zinc, cadmium, copper and lead ions in environmental samples using a thin-film mercury electrode and an artificial neural network. *Sensors (Basel)*. 2015;15(1):592-610.
doi: 10.3390/s150100592
78. Wang X, Lin W, Chen C, *et al.* Neural networks based fluorescence and electrochemistry dual-modal sensor for sensitive and precise detection of cadmium and lead simultaneously. *Sens Actuat B Chem*. 2022;366:131922.
doi: 10.1016/j.snb.2022.131922
79. Song K, Li H, Cheng G, *et al.* Learning Compact Discriminant Representation Via Low-Rank Bilinear Pooling. In: *IEEE Transactions on Pattern Analysis and Machine Intelligence*; 2025.
doi: 10.1109/TPAMI.2025.3601355
80. Ghosh S, Dissanayake K, Asokan S, Sun T, Rahman BMA, Grattan KTV. Lead (Pb2+) ion sensor development using optical fiber gratings and nanocomposite materials. *Sens Actuat B Chem*. 2022;364:131818.
doi: 10.1016/j.snb.2022.131818
81. Wilson D, Gutiérrez JM, Alegret S, Del Valle M. Simultaneous determination of Zn (II), Cu (II), Cd (II) and Pb (II) in soil samples employing an array of potentiometric sensors and an artificial neural network model. *Electroanalysis*. 2012;24(12):2249-2256.
doi: 10.1002/elan.201200440
82. Liu Y, Pu H, Sun DW. Efficient extraction of deep image features using convolutional neural network (CNN) for applications in detecting and analysing complex food matrices. *Trends Food Sci Technol*. 2021;113:193-204.
doi: 10.1016/j.tifs.2021.04.042
83. Guodong W, Lanxiang S, Wei W, Tong C, Meiting G. A feature selection method combined with ridge regression and recursive feature elimination in quantitative analysis of laser induced breakdown spectroscopy. *Plasma Sci Technol*. 2020;22(7):074002.
doi: 10.1088/2058-6272/ab76b4
84. Mantena S, Mahammood V, Rao KN. Prediction of soil salinity in the Upputeru river estuary catchment, India, using machine learning techniques. *Environ Monit Assess*. 2023;195(8):1006.
doi: 10.1007/s10661-023-11613-y
85. Meng H, Zhang J, Chang Y, Zheng Z. A new method for predicting chlorophyll-a concentration in a reservoir: Coupling EFDC hydrodynamic and water quality model with ConvLSTM-MLP network. *J Hydrol*. 2025;660:133485.
doi: 10.1016/j.jhydrol.2025.133485
86. Wang B, Liu T, Liu J, *et al*. A novel strategy of anti-interference for solid dilution in elemental analysis using spontaneous surface dispersion theory. *At Spectrosc*. 2020;41:119-126.
doi: 10.46770/AS.2020.03.004
87. Wang Y, Meng X, Shi W, *et al*. Single-Atom Cu anchored on a UiO-66 surface-enhanced raman scattering sensor for trace and rapid detection of volatile organic compounds. *Research*. 2025;8:0841.
doi: 10.34133/research.0841

88. Liu N, Ye W, Zhao G, Liu G. Development of smartphone-controlled and machine-learning-powered integrated equipment for automated detection of bioavailable heavy metals in soils. *J Hazard Mater*. 2024;465:133140. doi: 10.1016/j.jhazmat.2023.133140
89. Ye W, Liu N, Zhao G, Liu G. Accurate detection of Cd^{2+} and Pb^{2+} concentrations in soils by stripping voltammetry peak areas under the mutual interference of multiple heavy metals. *Metals*. 2023;13(2):270. doi: 10.3390/met13020270
90. Yao H, Wu R, Zou J, *et al.* A machine learning strategy-incorporated BiFeO3/Ti3C2 MXene electrochemical platform for simple, rapid detection of Pb^{2+} with high sensitivity. *Chemosphere*. 2023;340:139728. doi: 10.1016/j.chemosphere.2023.139728
91. Chakraborty TK, Rahman MS, Islam KR, *et al*. Application of machine learning and statistical approaches for optimization of heavy metals (Cd^{2+}, Pb^{2+}, Cu^{2+}, and Zn^{2+}) adsorption onto carbonized char prepared from PET plastic bottle waste. *AQUA Water Infrastructure Ecosyst Soc*. 2024;73(6):1097-1112. doi: 10.2166/aqua.2024.222
92. Aftab RA, Zaidi S, Danish M, Ansari KB, Danish M. Novel machine learning (ML) models for predicting the performance of multi-metal binding green adsorbent for the removal of Cd (II), Cu (II), Pb (II) and Zn (II) ions. *Environ Adv*. 2022;9:100256. doi: 10.1016/j.envadv.2022.100256
93. Ren Y, Yang W, Tan Z, Zhang L, Pan R. Highly sensitive detection of Pb^{2+} with a non-contact, near-infrared responsive hydrogel-functionalized optical fiber sensor. *J Hazard Mater*. 2024;480:136344. doi: 10.1016/j.jhazmat.2024.136344
94. Hegde RS. Deep learning: A new tool for photonic nanostructure design. *Nanoscale Adv*. 2020;2(3):1007-1023. doi: 10.1039/c9na00656g
95. Chugh S, Gulistan A, Ghosh S, Rahman BMA. Machine learning approach for computing optical properties of a photonic crystal fiber. *Opt Express*. 2019;27(25):36414-36425. doi: 10.1364/OE.27.036414
96. Bodelón G, Pastoriza-Santos I. Recent progress in surface-enhanced raman scattering for the detection of chemical contaminants in water. *Front Chem*. 2020;8:478. doi: 10.3389/fchem.2020.00478
97. Narayan A, Berger B, Cho H. Assessing single-cell transcriptomic variability through density-preserving data visualization. *Nat Biotechnol*. 2021;39(6):765-774. doi: 10.1038/s41587-020-00801-7
98. Tao Q, Liu T, Si M, Zhang T, Jin Q, Wang X. Ratiometric electrochemical aptasensor with magnetic separation-assisted entropy driven catalyst for dual detection of Pb^{2+} and Hg^{2+} in aquatic products: A high-amplification strategy for food safety monitoring. *J Environ Chem Eng*. 2025;13(5):118165. doi: 10.1016/j.jece.2025.118165
99. Bao Z, Al T, Couillard M, *et al*. A cross scale investigation of galena oxidation and controls on mobilization of lead in mine waste rock. *J Hazard Mater*. 2021;412:125130. doi: 10.1016/j.jhazmat.2021.125130
100. Abdel-Moneim AM, El-Toweissy MY, Ali AM, Awad Allah AAM, Darwish HS, Sadek IA. Curcumin ameliorates lead (Pb^{2+})-induced hemato-biochemical alterations and renal oxidative damage in a rat model. *Biol Trace Elem Res*. 2015;168(1):206-220. doi: 10.1007/s12011-015-0360-1
101. Braem DS, Parrott N, Hutchinson L, Steiert B. Introduction of an artificial neural network-based method for concentration-time predictions. *CPT Pharmacometrics Syst Pharmacol*. 2022;11(6):745-754. doi: 10.1002/psp4.12786
102. Yu J, Du M, Zhang Y, Chen X, Yang Z. Research progress on micro/nanopore flow behavior. *Molecules*. 2025;30(8):1807. doi: 10.3390/molecules30081807