ADBA
Intelligent Systems for Chemical and Materials Science
Cutting-Edge Scientific Solutions

e-ISSN:3150-0021
RESEARCH ARTICLE
Vol.1 / No.1 / 2026
pp. 21-25 / Article No. e004
https://doi.org/10.69882/adba.iscms.2026074

Advanced Spectral and Image-Based Detection of Rare Earth Elements Using Machine Learning

Ş. Samet Kaplan [1]
[1]Metallurgical and Materials Engineering Department, Faculty of Engineering and Natural Sciences, Hitit University, Çorum, Türkiye

ABSTRACT The escalating global demand for rare earth elements (REEs) in green energy and advanced technologies has made efficient resource assessment more critical than ever. However, conventional laboratory characterization methods in geological samples, such as inductively coupled plasma mass spectrometry and x-ray fluorescence, are often destructive, time-consuming, and cost-intensive. Consequently, to accelerate exploration and processing pipelines, there is an urgent need to transition toward rapid, in-situ, and non-destructive analytical techniques for mineralogical screening. This review explores the paradigm shift brought about by synergizing modern spectroscopic methods and microchemical imaging with advanced machine learning (ML) and data-driven architectures. Spectroscopic methods evaluated include visible to shortwave infrared spectroscopy, laser-induced breakdown spectroscopy, and mass spectrometry. Microchemical imaging focuses on scanning electron microscopy coupled with energy dispersive x-ray spectroscopy to map complex mineral phases. Traditional mineralogical and spectral analyses of complex geological ores and secondary resources (e.g., coal fly ash, e-waste, and bauxites) are frequently hindered by severe matrix effects, elemental interferences, and dense spectral overlaps. Here, it is highlighted how intelligent algorithmic models effectively decode these high-dimensional mineralogical datasets. Rather than focusing on pure computational architecture, this paper emphasizes how AI-driven frameworks autonomously extract hidden geochemical correlations, suppress geological background noise, and identify multi-elemental patterns or microstructural mineral inclusions. Ultimately, the deployment of ML-empowered optical sensors and automated mineral tracker algorithms significantly improves the detection limits, phase quantification, and prediction accuracy for REEs. This integration establishes a robust foundation for autonomous mineral exploration, efficient metallurgical resource screening, and the advancement of processing technologies within the circular economy.

KEYWORDS
Rare earth elements
Machine learning
Elemental characterization

INTRODUCTION

The demand for rare earth elements (REEs) is rapidly growing due to their irreplaceable applications in green energy, advanced electronics, and defense technologies (Balaram 2023). Currently, the global supply of REEs relies heavily on primary ore deposits, such as carbonatites, as well as unconventional secondary resources like karst bauxites, marine phosphorites, and coal combustion byproducts. Traditionally, the precise detection and quantification of REEs have depended on highly accurate laboratory techniques. For instance, in evaluating the heavy REE (HREE) distribution in mediterranean-type karst bauxites, researchers collected 111 samples from southern Italy, generating geochemical data via standard spectrometric extraction (Buccione *et al.* 2024). Similarly, the REE prospectivity of Tethyan phosphorites was modeled using 161 pure phosphorite samples collected from six different countries (Tahar-Belkacem *et al.* 2024). In coal environments, 91 Indiana coal samples were prepared following ASTM standards and analyzed to evaluate market-based outlook coefficients for REEs + yttrium

Manuscript received: 24 April 2026,
Revised: 28 May 2026,
Accepted: 20 July 2026,
Published: 16 August 2026.

[1]ssametkaplan@hitit.edu.tr (**Corresponding author**).

(REY) (Chatterjee *et al.* 2022).

While providing low detection limits, conventional bulk-analysis methods present significant operational bottlenecks: they are time-consuming, expensive, and require destructive sample preparation involving complex acid digestions (Balaram 2023; Zaremotlagh and Hezarkhani 2016). Consequently, there is an urgent need for in-situ, non-destructive optical sensors and imaging systems capable of providing real-time data directly in the field (Chen *et al.* 2020). Concurrently, machine learning (ML) models are now heavily utilized to process these high-dimensional datasets. By mapping nonlinear interactions such as leveraging and concentrations ML algorithms reliably predict REE enrichment and bypass the limitations of sparse data through augmentation techniques like synthetic minority over-sampling technique (SMOTE) (Chatterjee *et al.* 2022; Abedini *et al.* 2026).

SPECTRAL DETECTION OF REES AND MACHINE LEARNING INTEGRATION

The identification of REEs relies fundamentally on capturing their unique spectral signatures across different electromagnetic domains. However, traditional spectral analysis is frequently hindered by severe matrix effects, dense emission lines, and background noise. Modern analytical frameworks utilize advanced machine learning (ML) models to decode these high-dimensional datasets, directly detecting and quantifying REEs presence from reflectance, absorbance, optical emission, and mass spectra:

Reflectance and Absorbance Spectra: The identification of specific mineralogical phases often relies on analyzing how materials interact with light. One of the prominent methods used for evaluating reflectance and absorbance spectra is visible to shortwave infrared (VNIR-SWIR) spectroscopy. It is a passive optical method based on the selective absorption and reflectance of electromagnetic radiation across the 400 to 2500 nm wavelength range without destroying the sample. It operates by capturing the properties of light across a set of continuous, narrow spectral bands to generate complex "data cubes," where each image pixel contains a characteristic, continuous spectrum. In geological environments, the spectral absorption features detected in the VNIR region typically arise from transitional elements, including iron-bearing minerals and REEs, whereas the SWIR region is primarily utilized for identifying associated alteration mineral assemblages, such as clays, carbonates, and phyllosilicates (Koerting *et al.* 2024). For REEs specifically, the utility of this technique stems from the unique 4f electron transitions of rare earth ions. However, extracting these signatures is challenging; for instance, prominent absorption features at 580, 744, 804, and 872 nm are often severely obscured by the overlapping absorptions of iron oxides like goethite. To resolve this spectral interference, advanced ML algorithms are deployed. Murer and de Souza Filho (2026) integrated hyperspectral drill core scans from the Nolans Bore deposit with partial least squares regression (PLSR), Light gradient-boosting machine (LightGBM), and multiple-instance learning (MIL) neural networks. While PLSR provided linear physical interpretability, LightGBM successfully captured complex nonlinear interactions involving gangue minerals. Notably, the attention-based MIL neural network, operating at a "bag-level" to aggregate spectral diversity, effectively decoupled the REE absorption depths from the noisy iron oxide background. This deep learning approach achieved a high global prediction accuracy for total REE (TREE) grades with an of 0.86 and a root mean square error (RMSE) of 0.97, demonstrating that ML can translate passive optical data into highly accurate geochemical grades (Murer and de Souza Filho 2026). On a broader scale, drone and satellite-mounted VNIR-SWIR hyperspectral imaging combined with classification algorithms successfully detected Nd anomalies in open-field complexes at remarkably low detection limits <200 (Balaram 2023).

Optical Emission Spectra: To evaluate elemental fingerprints through light emission, researchers rely on specialized active techniques. One of the highly versatile methods utilizing optical emission spectra is inductively coupled plasma optical emission spectroscopy (ICP-OES). It operates by introducing liquid samples into an inductively coupled argon plasma sustained at extremely high temperatures (8000–10000 K). Within this plasma, the sample undergoes desolvation, vaporization, atomization, and excitation. As the excited atoms and ions return to lower energy states, they emit photons at characteristic wavelengths, which are subsequently separated by a monochromator or polychromator and measured by a detector to quantify elemental abundance (Khan *et al.* 2022). Another prominent method in this category is laser-induced breakdown spectroscopy (LIBS). It is an active, micro-destructive analytical technique that utilizes a high-energy pulsed laser focused directly on the sample surface to generate a localized, high-temperature plasma. As the plasma cools, the excited atoms and ions revert to their ground states, emitting light at characteristic wavelengths (Chen *et al.* 2020). However, the dense emission lines of REEs in complex matrices cause severe spectral overlaps, making standard univariate analysis fail. Deep learning architectures revolutionized this domain by decoding complex spectral sequences. Yang *et al.* (2025) developed an advanced ensemble model (iTBi-RF-LIBS) combining transformer attention mechanisms, bidirectional long short-term memory (BiLSTM), and random forest algorithms to process LIBS spectra of rare earth ores. The attention mechanism assigns specific weights to different wavelengths, capturing inter-wavelength correlations and suppressing local noise, while the BiLSTM structure models sequential dependencies. This AI-driven approach successfully isolated target ionic emissions (e.g., La at 333.72 nm, Ce at 412.72 nm) from severe background noise, achieving exceptional calibration coefficients and reducing the root mean square error of prediction to 0.069–0.184 wt% (Yang *et al.* 2025). Similarly, in the context of e-waste recycling, Tian *et al.* (2023) utilized unsupervised principal component analysis (PCA) to reduce the dimensionality of dense emission spectra, coupled with a supervised backpropagation artificial neural network (BP-ANN). This ML framework effectively differentiated the emission spectra of Gd, Y, and Yb in electronic waste, achieving an exceptional recognition accuracy of 99.9% (Tian *et al.* 2023).

Mass Spectra: For high-throughput and ultra-sensitive elemental quantification, assessing the mass-to-charge ratio of ionized particles is essential. One of the primary methods used to investigate mass spectra is inductively coupled plasma mass spectrometry (ICP-MS). It relies on introducing a liquid sample via a nebulizer into an ultra-hot argon plasma, where the sample is evaporated, dissociated, atomized, and ionized. The resulting singly charged positive ions are then extracted through a differentially pumped interface into a mass spectrometer (such as a quadrupole), which separates and quantifies the ions based on their specific mass-to-charge ratios (Longerich *et al.* 1990). While standard ICP-MS requires the sample to be completely dissolved into a liquid solution via complex acid digestion, an alternative approach is Laser Ablation ICP-MS (LA-ICP-MS). It bypasses the dissolution bottleneck by utilizing a pulsed laser to directly ablate a solid sample, generating an aerosol of micro-particles that is transported by an inert carrier gas into the ICP for ionization and subsequent mass analysis, allowing for spatially resolved

micro-destructive sampling (Mokgalaka and Gardea-Torresdey 2006). Processing the massive datasets generated by mass spectrometry requires robust ML frameworks to mitigate noise and data sparsity. Song *et al.* (2024) demonstrated the power of deep learning in this domain by training a multi-task neural network on ICP-MS data from 99 diverse coal and coal ash samples. Unlike traditional single-task models that predict one element at a time, this multi-task neural network simultaneously predicts the concentrations of multiple individual REEs. This synergistic approach provides implicit data augmentation and regularization, exploiting the internal geochemical correlations among REEs to filter out individual measurement noise. Consequently, the multi-task model achieved a superior test set of 0.89, significantly outperforming single-task regressions and providing a highly robust tool for rapid REE screening in secondary resources (Song *et al.* 2024). At the micro-scale, LA-ICP-MS bypasses complex acid dissolution to provide extremely high-resolution 2D and 3D spatial mapping of critical metals. The massive isotopic data cubes generated by these direct solid analyses increasingly rely on algorithmic processing to map REEs directly within deep-ocean manganese nodules and individual detrital grains (Balaram 2023).

IMAGE-BASED DETECTION OF REES AND MACHINE LEARNING INTEGRATION

Beyond spectral signals, advanced image processing and computer vision extend REE characterization into intelligent 2D/3D spatial, morphological, and topological recognition. However, manual interpretation of multi-elemental maps and complex microstructures is highly subjective, inefficient, and prone to error when dealing with highly heterogeneous materials. To automate and optimize this process, researchers increasingly rely on advanced machine learning algorithms to identify REE-bearing phases directly from microscopic visualizations.

Extracting morphological topography alongside high-resolution elemental distribution maps is critical for understanding heterogeneous secondary resources. One of the most effective methods for obtaining microchemical images is scanning electron microscopy coupled with energy dispersive x-ray spectroscopy (SEM-EDS). This technique operates by scanning a sample surface with a focused electron beam to generate high-resolution morphological images and corresponding elemental maps simultaneously. However, deciphering the massive amount of overlapping compositional data generated by SEM-EDS requires robust computational tools. To address this, Antipovych (2026) applied unsupervised machine learning to automate the characterization of coal fly ash. By extracting standardized high-dimensional SEM-EDS data specifically Al, Si, Fe, and Ti mass fractions combined with particle diameter the researcher deployed a k-means clustering algorithm to identify hidden patterns. Evaluated and confirmed through principal component analysis (PCA) scatter plots and ANOVA tests, the AI model successfully segregated the aluminosilicate cenospheres into three distinct chemical-morphological subtypes (Ti-rich/Si-poor, Fe-Si-rich, and Al-rich/Fe-poor). Crucially, the algorithm autonomously pinpointed the Al-rich cluster as the primary host for REEs, providing a data-driven basis for targeted beneficiation strategies without the need for destructive bulk dissolution (Antipovych 2026). Similarly, multimodal image analysis frameworks have been combined with data science methods to autonomously isolate trace REEs mineral phases directly within complex coal utilization byproducts, significantly accelerating resource assessment (Montross *et al.* 2018; Creason *et al.* 2023).

Recognizing complex structural inclusions within microstructures also requires advanced pattern-recognition algorithms like convolutional neural networks (CNNs). It uses convolutional layers to automatically extract hierarchical visual features (like edges, textures, and specific shapes) from raw input grids, eliminating the need for manual feature engineering. For instance, Ma and Huang (2021) utilized computer vision to detect trace inclusions of rare earth elements, such as cerium (Ce) and lanthanum (La). Although the title of their study implies application to rocks and minerals, their actual experimental focus was on identifying inclusions within 200/300 series stainless steel alloys (e.g., 304-grade steel) used for offshore platforms. The location of these trace elements whether segregated at grain boundaries or forming sulfide inclusions directly impacts the material's mechanical strength and corrosion resistance. Since traditional target detection algorithms were slow and computationally inadequate, the researchers optimized a faster R-CNN target detection network by fusing it with VGG16 and MobileNet classification architectures. After applying various industrial heat treatments to the steel samples, such as normalizing, subcritical quenching, and tempering, the trained deep learning model was deployed on the resulting complex microstructural images. The AI system successfully and autonomously identified cerium-sulfide phases, improving trace element detection accuracy by more than 50% compared to classical methods (Ma and Huang 2021). While highly effective for metallurgical quality control and alloy engineering, applying these specific computer vision models to raw, heterogeneous geological ores would require careful adaptation to account for vastly different textural and spectral complexities.

Summary of Applied AI Methods and Results: Table 1 provides a comparative summary of the machine learning methods discussed, their specific applications, and their predictive performance metrics based on the reviewed studies.

CONCLUSION

The integration of direct spectral detection (reflectance, emission, and mass spectra) and multimodal imaging with deep learning represents a significant paradigm shift, transitioning the industry from slow, destructive laboratory analyses toward rapid, autonomous sensing. Implementing these technologies requires carefully balancing model trade-offs. Deep learning structures like transformers and CNNs excel at resolving complex matrix interferences such as isolating specific Nd/Ce emission lines, unmixing iron oxide spectral overlaps, or identifying microstructural inclusions yet they demand substantial training data and often lack the physical interpretability provided by simpler linear models. Strategies to mitigate data sparsity include data augmentation techniques like SMOTE and the deployment of multi-task neural networks to simultaneously predict multiple REE concentrations with enhanced noise resistance. Ultimately, deploying these AI-empowered optical sensors and image-processing algorithms directly supports the circular economy—facilitating e-waste sorting and screening reclaimed coal ashes—and the future of autonomous mineral exploration in complex geological environments.

To bridge the critical gap between computational predictions and real-world industrial mineral processing, several core future research avenues must be pursued. Purely data-driven architectures often predict statistically sound but physically rare mineral relationships due to severe matrix overlaps and chemical interferences. Therefore, future analytical frameworks must incorporate rigid thermodynamic, crystallographic, and geochemical constraints directly into algorithmic loss functions to suppress artificial back-

Table 1 Summary of machine learning methods used for elemental detection.

Reference Study	Characterization Method	Analytical Matrix	Target Elements	Algorithmic Model	R^2 value
Abedini et al.	XRF, ICP-MS	Karst Bauxite	Total REEs	XGBoost	0.934 (LREE) 0.770 (HREE)
Murer & de Souza Filho	VNIR-SWIR Spectroscopy	Fluorapatite Veins	Total REEs	Partial Least Squares (PLS)	0.86
Yang et al.	LIBS	Rare Earth Ores	La, Ce, Nd	Ensemble (iTBi-RF-LIBS)	0.996
				Deep Learning (BiLSTM)	0.992
Song et al.	LA-ICP-MS	Karst Bauxite	Lanthanides	Supervised Regression	0.89
Tahar-Belkacem et al.	ICP-AES, XRF & ICP-MS	REEs Phosphorites	Light and heavy REEs	XGBoost	0.979 (LREE) 0.961 (HREE)
Koerting et al.	VNIR-SWIR Imaging	Iron oxide matrices	Nd, Sm, Dy	Convolutional Neural Net	0.914

ground noise and eliminate outlier predictions. Furthermore, a major bottleneck in deployed machine learning models is their intense dependency on massive, site-specific calibration datasets. Models trained on well-documented primary carbonatite deposits often fail when deployed on highly heterogeneous secondary resources like coal fly ash, e-waste, or weathered laterites. To resolve this domain vulnerability and maximize the industrial utility of automated mineralogical screening, subsequent studies should focus on the following key operational pathways:

- **Development of Physics-Informed Neural Networks (PINNs):** Embedding fundamental mineralogical rules such as strict stoichiometric limits of phosphate-bearing phases (e.g., monazite, apatite) or ionic substitution constants within karst bauxites directly into the machine learning training loop to ensure physically valid geochemical predictions.
- **Cross-Domain Transfer Learning Optimization:** Utilizing unsupervised domain adaptation techniques to design foundational models that are pre-trained on standardized synthetic mineral spectra and can successfully generalize to unmapped geological matrices with minimal target site calibration data.
- **Flowsheet Integration with Downstream Hydrometallurgy:** Linking sensor-based bulk ore sorting systems directly with chemical recovery operations by using real-time mineralogical variations, matrix compositions, and liberation sizes to dynamically predict reagent consumption, dissolution kinetics, and overall extraction efficiencies within the recovery plant.

Ethical Standard

The author have no relevant financial or non-financial interests to disclose.

Funding

This research received no external funding.

Data Availability Statement

The data supporting the findings of this study are available from the corresponding author upon reasonable request.

Conflict of Interest

The author declares no conflict of interest.

LITERATURE CITED

Abedini, A., M. Khosravi, O. Ameur-Zaimeche, R. Kechiched, A. Ouladmansour, *et al.*, 2026 A machine learning framework to lanthanide element distribution and predictability from the northwestern Iranian karst bauxite deposits. Earth Science Informatics **19**: 30, https://doi.org/10.1007/s12145-026-02094-x.

Antipovych, Y., 2026 AI-assisted identification of REE-bearing cenospheres in coal fly ash. The European Physical Journal Special Topics https://doi.org/10.1140/epjs/s11734-026-02180-7.

Balaram, V., 2023 Advances in analytical techniques and applications in exploration, mining, extraction, and metallurgical studies of rare earth elements. Minerals **13**: 1031, https://doi.org/10.3390/min13081031.

Buccione, R., O. Ameur-Zaimeche, A. Ouladmansour, R. Kechiched, and G. Mongelli, 2024 Data-centric approach for predicting critical metals distribution: Heavy rare earth elements in cretaceous Mediterranean-type karst bauxite deposits, southern Italy. Geochemistry **84**: 126026, https://doi.org/10.1016/j.chemer.2023.126026.

Chatterjee, S., M. Mastalerz, A. Drobniak, and C. Ö. Karacan, 2022 Machine learning and data augmentation approach for identification of rare earth element potential in Indiana coals, USA. International Journal of Coal Geology **259**: 104054, https://doi.org/10.1016/j.coal.2022.104054.

Chen, T., T. Zhang, and H. Li, 2020 Applications of laser-induced breakdown spectroscopy (LIBS) combined with machine learning in geochemical and environmental resources exploration. Trends in Analytical Chemistry **133**: 116113, https://doi.org/10.1016/j.trac.2020.116113.

Creason, C. G., D. M. Justman, M. K. Mark-Moser, R. B. Thomas, A. Bean, *et al.*, 2023 A geo-data science method for assessing unconventional rare-earth element resources in sedimentary systems. Natural Resources Research **32**, https://doi.org/10.1007/s11053-023-10163-x.

Khan, S. R., B. Sharma, P. A. Chawla, and R. Bhatia, 2022 Inductively coupled plasma optical emission spectrometry (ICP-OES): A powerful analytical technique for elemental analysis. Food Analytical Methods **15**: 666–688, https://doi.org/10.1007/s12161-021-02148-4.

Koerting, F., S. Asadzadeh, J. C. Hildebrand, E. Savinova, E. Kouzeli, *et al.*, 2024 VNIR-SWIR imaging spectroscopy for mining: Insights for hyperspectral drone applications. Mining **4**: 1013–1057, https://doi.org/10.3390/mining4040057.

Longerich, H. P., G. A. Jenner, B. J. Fryer, and S. E. Jackson, 1990 Inductively coupled plasma-mass spectrometric analysis of geological samples: A critical evaluation based on case studies. Chemical Geology **83**: 105–118, https://doi.org/10.1016/0009-2541(90)90143-U.

Ma, S. and W. Huang, 2021 Application of deep learning algorithms in determination of trace rare earth elements of cerium group in rocks and minerals. Wireless Communications and Mobile Computing **2021**: 9945141, https://doi.org/10.1155/2021/9945141.

Mokgalaka, N. S. and J. L. Gardea-Torresdey, 2006 Laser ablation inductively coupled plasma mass spectrometry: Principles and applications. Applied Spectroscopy Reviews **41**: 131–150, https:

//doi.org/10.1080/05704920500510703.
Montross, S. N., C. A. Verba, H. L. Chan, and C. Lopano, 2018 Advanced characterization of rare earth element minerals in coal utilization byproducts using multimodal image analysis. International Journal of Coal Geology **195**: 362–372, https://doi.org/10.1016/j.coal.2018.06.018.
Murer, T. C. and C. R. de Souza Filho, 2026 VNIR-SWIR spectroscopy and machine learning: A new approach for rare earth element detection and quantification in fluorapatite veins (Nolans Bore deposit, Australia). Journal of Geochemical Exploration **282**: 107955, https://doi.org/10.1016/j.gexplo.2025.107955.
Song, Y., Y. Zhao, A. Ginella, B. Gallagher, G. Sant, *et al.*, 2024 Predicting rare earth elements concentration in coal ashes with multi-task neural networks. Materials Horizons **11**: 1448–1464, https://doi.org/10.1039/d3mh01491f.
Tahar-Belkacem, N., O. Ameur-Zaimeche, R. Kechiched, A. Ouladmansour, S. Heddam, *et al.*, 2024 Machine learning models to predict rare earth elements distribution in Tethyan phosphate ore deposits: Geochemical and depositional environment implications. Geochemistry **84**: 126128, https://doi.org/10.1016/j.chemer.2024.126128.
Tian, D., L. Tian, L. Shen, Y. Ge, Z. Sun, *et al.*, 2023 Rare earth metals detection and recognition based on laser induced breakdown spectroscopy and machine learning. Optics Express **31**: 20545–20558, https://doi.org/10.1364/OE.493905.
Yang, J., S. Li, Z. Zhang, D. Qian, S. Sun, *et al.*, 2025 Transformer-based deep learning models for quantification of La, Ce, and Nd in rare earth ores using laser-induced breakdown spectroscopy. Talanta **292**: 127937, https://doi.org/10.1016/j.talanta.2025.127937.
Zaremotlagh, S. and A. Hezarkhani, 2016 A geochemical modeling to predict the different concentrations of REE and their hidden patterns using several supervised learning methods: Choghart iron deposit, Bafq, Iran. Journal of Geochemical Exploration **165**: 35–48, https://doi.org/10.1016/j.gexplo.2016.02.001.

How to cite this article: Kaplan, Ş. S. Advanced Spectral and Image-Based Detection of Rare Earth Elements Using Machine Learning. *Intelligent Systems for Chemical and Materials Science*, 1(1), 21-25, 2026.

Licensing Policy: The published articles in ISCMS are licensed under a Creative Commons Attribution-NonCommercial 4.0 International License.

