BULLETIN
OF "CAROL I" NATIONAL DEFENCE UNIVERSITY

BULLETIN
OF "CAROL I" NATIONAL DEFENCE UNIVERSITY

https://buletinul.unap.ro/index.php/en/

Strengthening Nuclear Security with ML: Full-Spectrum ^{137}Cs Burial Depth Estimation

Konstantinos KARAFASOULIS, PhD*

*Laboratory Teaching Staff, Hellenic Army Academy, Athens, Greece
e-mail: ckaraf@gmail.com

Abstract

The non-intrusive characterization of buried radioactive sources is a critical capability for thwarting illicit trafficking, mitigating orphan-source hazards, and safeguarding civilian populations against radiological threats. Depth estimation, in particular, enables rapid threat assessment and informed countermeasure deployment following incidents such as transnational uranium diversion or the loss of medical and industrial sources. In this feasibility study, we demonstrate a machine learning approach to estimate the burial depth of a ^{137}Cs point source in dry sand over the range of 5–95 cm. Our method employs gradient-boosted decision trees trained on simulated full gamma-ray spectra partitioned into 1024 energy bins, thereby exploiting subtle variations across both the Compton continuum and multiple photopeaks. After hyperparameter tuning, the model achieved an average depth-estimation standard deviation of 5 cm across the full depth range. By leveraging the entire spectral profile rather than isolated peak ratios, this algorithm delivers enhanced accuracy and robustness in heterogeneous field conditions. The results validate the potential of full-spectrum, gradient boosted models as field-deployable tools for rapid subsurface threat localization, reinforcing layers of nuclear security and environmental monitoring efforts worldwide.

Keywords:

Nuclear Security; CBRN; Artificial Intelligence; Gradient Boosted Decision Trees; Buried Radiation Sources; Gamma Radiation.

Article info

Received: 1 August 2025; Revised: 23 October 2025; Accepted: 14 November 2025; Available online: 15 January 2026

Citation: Karafasoulis, K. 2025. "Strengthening Nuclear Security with ML: Full-Spectrum 137Cs Burial Depth Estimation." *Bulletin of "Carol I" National Defence University*, 14(4): 53-64. https://doi.org/10.53477/2284-9378-25-58

Introduction

Non-intrusive localization and depth characterization of subsurface radioactive materials sit at the nexus of global security and public safety. In January 2025, Takeshi Ebisawa—a senior Yakuza leader—pleaded guilty in Manhattan to conspiring to traffic uranium and weapons-grade plutonium from Myanmar, allegedly intending to supply Iran's clandestine weapons program (U.S. Department of Justice 2024). Just six months later, on July 17, 2025, Georgian security forces detained two individuals in Batumi for attempting to sell $3 million worth of uranium—enough to build a rudimentary nuclear device—underscoring the persistent vulnerability of fissile materials to illicit diversion (Reuters 2025).

Beyond state-level trafficking, the proliferation of "orphan" sources poses an equally grave risk to communities worldwide. In December 2013, a 111 TBq cobalt-60 teletherapy source was hijacked en route from Tijuana to a Mexican waste facility; thieves removed its shielding and abandoned it in a field near Hueypoxtla, creating a severe radiological hazard before recovery (IAEA 2013). In January 2023, a caesium-137 capsule, used as a nucleonic level sensor in a Western Australia mine, detached from a truck and remained missing for nearly three weeks, prompting a statewide radiation alert and demonstrating how easily high-activity sources can become lost in transit (ABC News 2023). These examples highlight the urgent need for tools that can not only detect radiation but also estimate its depth below ground to inform safe recovery and containment actions.

This need is especially critical in Chemical, Biological, Radiological, and Nuclear (CBRN) defense contexts, where rapid, non-intrusive characterization of contaminated environments is essential for safe and effective response. In radiological release scenarios or during interdiction of illicit materials, first responders and CBRN teams must determine not just the presence of radioactive sources but also their subsurface positioning to guide protective measures and mitigation strategies. Conventional handheld detectors may signal increased dose rates, but without depth estimation, responders risk improper triage or unsafe excavation. The method presented in this work addresses this operational gap by delivering real-time, data-driven estimates of source burial depth using machine learning on full-spectrum gamma-ray data. This provides actionable intelligence for CBRN missions, enhancing both personnel safety and decision-making efficiency in complex field environments.

In current practice, responders can detect elevated dose rates but lack rapid, accurate depth estimation to guide standoff procedures and excavation planning. Traditional depth-estimation techniques typically analyze the ratio of Compton-continuum counts to photopeak counts in a gamma-ray spectrum, leveraging the fact that deeper burial alters the relative intensities of these spectral features. While effective in controlled settings, peak-ratio methods can suffer from reduced accuracy in heterogeneous soils, variable moisture conditions, and low-count regimes common in real-world deployments.

To target this gap, this paper presents a novel algorithm based on gradient boosted decision trees that harnesses the entire gamma-ray spectrum—divided into 1024 energy bins—as input features. By exploiting subtle variations across the full spectral profile rather than focusing solely on isolated peaks, our approach delivers enhanced depth-estimation accuracy and robustness across diverse soil types and isotope mixtures. This comprehensive spectral machine-learning framework represents a significant advance for global security applications, enabling faster threat interdiction, more reliable site characterization, and improved protection of both national borders and civilian populations.

Related Work

Non-intrusive depth estimation of buried radioactive sources has evolved through a succession of increasingly sophisticated techniques. Early analytical models assumed homogeneous media and fit simple attenuation laws to photopeak or scatter-continuum counts. For example, (Ukaegbu and Gamage 2018) introduced a novel 3D linear attenuation method—validated in sand up to 12 cm depth—by fitting logarithmic count-rate ratios across a spatial grid to infer ^{137}Cs burial depth. That year, they also proposed an extension using a CZT detector, demonstrating accurate depth estimation for ^{137}Cs and ^{60}Co sources buried at greater depths in sand (Ukaegbu and Gamage 2018).

To address soil heterogeneity, (Ukaegbu, Aspinall and Gamage 2019) combined ground-penetrating radar (GPR)–derived bulk density estimates with gamma attenuation curves, achieving reliable ^{137}Cs depth estimation up to ~25 cm in mixed soils. Parallel work by (Shippen and Joyce 2010) implemented a dual-photopeak ratio method for ^{137}Cs in concrete, using a NaI detector to estimate the depth of ^{137}Cs up to ~15mm.

More recently, machine learning techniques have shown increasing promise in gamma-ray spectroscopy by enabling automated analysis and pattern recognition in complex spectral environments. In their comparative study, (Kamuda, Zhao and Huff 2020) evaluated artificial neural networks (ANNs) and convolutional neural networks (CNNs) for the identification of radioisotope mixtures, demonstrating that these models can effectively replicate the decision-making process of trained spectroscopists—even under challenging conditions such as calibration drift and unknown background radiation fields. Their findings highlight the potential of deep learning models to operate reliably in diverse gamma spectroscopy scenarios. However, most existing applications have focused on isotope identification rather than continuous parameter regression, and often rely on discrete classification or mixture recognition rather than detailed physical characterization, such as burial depth estimation.

Building on this body of work, our approach is to apply gradient boosted decision trees directly to the entire 1024-bin gamma spectrum—sans manual feature extraction—for continuous depth regression over 5–95 cm, leveraging a KNIME-based workflow and Geant4 simulations to deliver high accuracy and robustness across realistic background conditions.

Methodology

This study comprises two main components: (1) generation of a controlled, physics-based spectral dataset via Geant4 simulation, and (2) formulation and optimization of a gradient boosted decision tree (GBDT) regression model to predict burial depth from the full gamma-ray spectrum. Below, we describe each step in detail.

1.1. Geant4 Simulation Framework

As shown in Figure 1, all spectra were generated in a controlled Geant4 (Agostinelli, et al. 2003) simulation environment. A ^{137}Cs point source (662 keV) was positioned at depths from 5 cm to 95 cm measured from the sand surface, within a cubical block of dry sand measuring 1 m × 1 m × 1 m (ρ = 1.60 g/cm³, SiO_2 composition). A 3″ × 3″ NaI(Tl) scintillation detector was centered directly above the source axis and fixed at a height of 35 cm above the sand top. Photon transport processes—including photoelectric absorption, Compton scattering, and Rayleigh scattering—were simulated using Geant4.10 with a combined physics list incorporating both QGSP_BIC_HP for accurate hadronic interactions and PENELOPE for detailed low-energy electromagnetic processes. This configuration ensured high-fidelity modeling of gamma-ray interactions in the detector and surrounding materials across a broad energy range. The energy deposited in the NaI(Tl) detector volume was recorded in 1024 uniform channels spanning 0–1.6 MeV. These raw Monte Carlo spectra, generated under consistent geometry and physics settings, served as the foundation for all subsequent data augmentation and machine learning analysis.

TABLE 1. Simulation Parameters

Parameter	Value
Isotope	^{13}7Cs (662 keV gamma)
Medium	Dry sand (ρ = 1.60 g/cm³; SiO_2 composition)
Burial depths	5 cm – 95 cm, in 5 cm increments (19 settings)
Detector	NaI(Tl) scintillator, 3″ × 3″
Detector height above surface	35 cm
Energy range/bins	0 – 1.6 MeV, 1024 uniform bins
Generated events per depth	~4×109

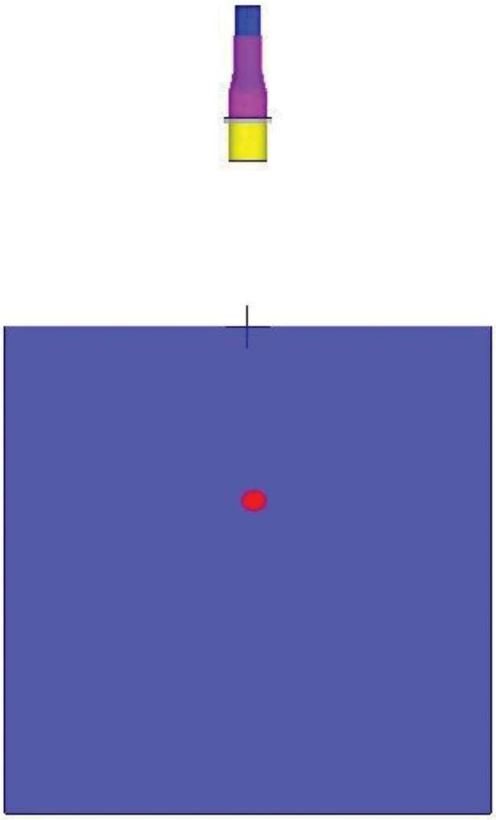

Figure 1 Geant4 simulation setup: a 3 in × 3 in NaI(Tl) detector placed 35cm above a 1m³ sand block, with a 137Cs source buried at depths of 5–95 cm; spectra recorded in 1024 channels (0–1.6 MeV).

1.2. Gamma Spectrum Dataset Generation

For each of the 19 burial depths, a high-statistics Geant4 simulation was performed in which approximately 4×10^9 gamma photons were emitted from a ^{137}Cs point source, capturing all relevant photon interactions in the detector. From this sample, we generated 500 spectra by bootstrap sampling with replacement. For each spectrum, the total number of sampled counts, $N_{\text{sig}}N_{\text{sig}}$, Was drawn uniformly between 5000 and 100000, thereby emulating variations in acquisition time or source activity. Each sampled event energy E was Gaussian smeared to reflect detector resolution with a full width at half maximum (FWHM) as in eq (1):

$$FWHM = A\sqrt{\frac{E}{E_c}} \tag{1}$$

where A = 19.2 keV and E_c = 662 keV. The smeared energies were then histogrammed into the 1024-channel spectrum.

A signal-fraction coefficient, α, was then sampled uniformly from 0.1 to 1.0, defining the desired proportion of signal in the final spectrum. From these values, the required number of background events was calculated as

$$N_{\mathrm{bg}} = N_{\mathrm{sig}} \frac{1-\alpha}{\alpha} \quad (2)$$

To assemble the background contribution, N_{bg} events were then randomly drawn (with replacement) from the measured background event list—i.e., the recorded event energies comprising the background spectrum. These sampled background energies were binned into the same 1024-channel histogram as the signal.

Finally, the smeared-signal histogram $S_{\mathrm{signal}}[i]$ (containing N_{sig} counts) and the sampled-background histogram $S_{\mathrm{bg}}[i]$ (containing N_{sig} **counts) were combined channel-wise:**

$$S_{\mathrm{final}}[i] \; = \; S_{\mathrm{signal}}[i] \; + \; S_{\mathrm{bg}}[i] \quad (3)$$

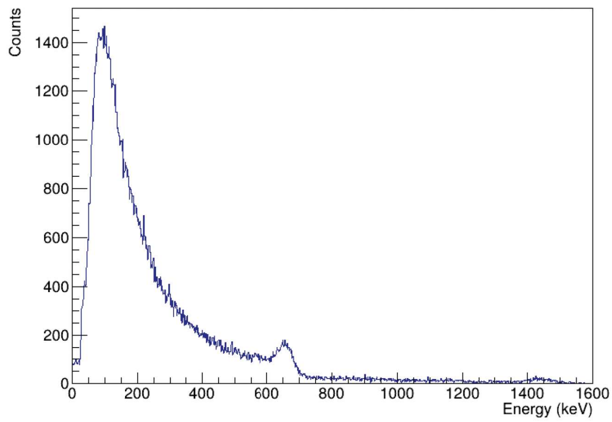

Figure 2 Energy spectrum histogram for a 137Cs source buried 30cm below the sand surface, measured by a 3″ × 3″ NaI(Tl) detector positioned 35cm above the ground. The x-axis shows deposited energy (keV) and the y-axis shows counts per channel, with the prominent 662 clearly visible

By construction, the total counts in S_{final} equal $N_{\mathrm{sig}} + N_{\mathrm{bg}}$, and exactly fraction α of those events originate from the ^{137}Cs signal, while the remainder reflects the authentic, measured background. Thus, the resulting dataset inherently spans both weak and strong source conditions as well as low and high background scenarios. By this approach, the training set possesses the advantages of broad count-statistical variation and diverse signal-to-noise ratios, "stressing" the gradient-boosted decision-tree model so that it learns depth-sensitive spectral features robustly across the full range of realistic nuclear security environments. This procedure yielded a total of 9500 full-spectrum training examples (as in Figure 2) that encapsulate realistic count fluctuations, detector resolution effects, and environmental background, ready for input into the regression model.

1.3. Gradient Boosted Decision Trees for Depth Regression

To translate the rich, high-dimensional spectral information into accurate depth estimates, we adopted gradient boosted decision trees (GBDT) (Friedman 2001) within the KNIME Analytics Platform (Berthold, et al. 2007), an ensemble learning method renowned for its ability to capture complex, nonlinear relationships without extensive manual feature engineering. Gradient boosted decision trees build a series of simple decision trees in sequence, where each new tree focuses on correcting the errors of the previous ensemble. Rather than fitting one large, complex model, GBDT constructs many small "weak learners" and combines their outputs to form a strong overall predictor. This strategy excels when the input space is high-dimensional—as is the case with 1024-bin gamma spectra—because the model can automatically identify and exploit subtle interactions among features (e.g., fine shifts in Compton continuum shape or slight photopeak distortions that correlate with burial depth).

Hyperparameter tuning was carried out via a structured optimization loop on a held-out test set (30 % of the total data). We explored a two-dimensional grid over:

- **Maximum tree depth:** from 3 to 10 levels, controlling how many successive splits each tree can make.
- **Number of trees (estimators):** from 70 to 200 total models in the ensemble.

The learning rate was fixed at 0.1 based on preliminary trials, indicating a good balance between convergence speed and generalization. For each candidate configuration, the model was trained on the 70 % training split and evaluated on the 30 % test split, using the minimization of the root-mean-square error (RMSE) as the selection criterion (eq. 4):

$$RMSE = \sqrt{\frac{1}{N}\sum_{i=1}^{N}\left(d_i^{\text{pred}} - d_i^{\text{true}}\right)^2} \tag{4}$$

Where:

- d_i^{true} is the known depth (in cm) of the $i-th$ spectrum,
- d_i^{pred} is the depth estimated by the GBDT model, and
- N is the total number of spectra in the test set

The optimal settings—seven layers per tree and 200 trees—were chosen because they yielded the lowest RMSE on the hold-out dataset.

By combining full-spectrum input with a finely tuned GBDT ensemble, our approach leverages both broad continuum shapes and narrow peak features, yielding a regression model capable of inferring ^{137}Cs burial depths with high accuracy and robustness across the 5 – 95 cm range. The trained model evaluates a single 1024-bin spectrum well under one second on commodity laptop hardware, enabling in-field, near-real-time depth estimates.

https://doi.org/10.53477/2284-9378-25-58

Results

The GBDT model, trained on a comprehensive dataset of Geant4-simulated gamma-ray spectra, demonstrated strong performance in estimating the burial depth of a ^{137}Cs radioactive source under diverse and realistic conditions. The dataset used for testing included spectra generated with a wide range of source strengths (total signal counts between 5,000 and 100,000) and background conditions (signal-to-total ratios, SNR, between 0.1 and 1.0), ensuring a robust evaluation across scenarios relevant to nuclear security operations.

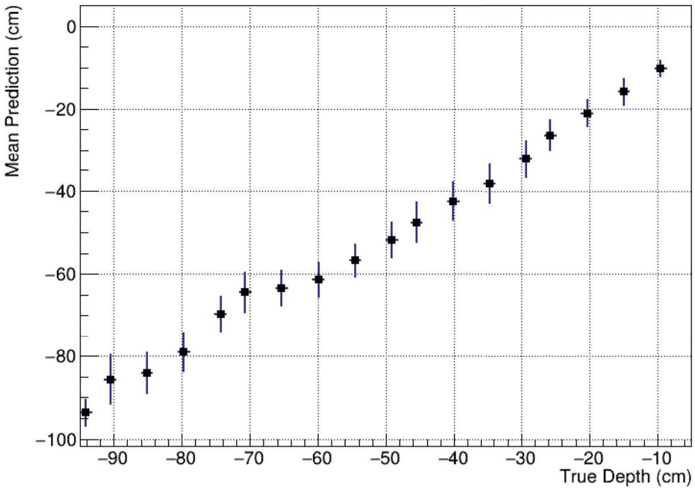

Figure 3 Mean predicted depth versus true burial depth for the GBDT model across the full range of 5–95 cm. Error bars represent the standard deviation of predictions at each depth point, reflecting the model's precision. The close alignment with the 1:1 line and consistent ~5 cm standard deviation demonstrate that the model maintains high accuracy and stability across all tested depths, regardless of signal-to-noise ratio.

Quantitatively, the model achieved a coefficient of determination of $R^2 = 0.96$, indicating that 96% of the variance in source depth was captured by the spectral features learned from the training data. The RMSE of the predictions was approximately 5cm across the full depth range of 5 to 95cm, demonstrating a high degree of accuracy even in the presence of varying noise levels and statistical fluctuations. As illustrated in Figure 3, the mean predicted depth tracks very closely with the true burial depth across the entire range. Importantly, the standard deviation of the predictions remains consistently around 5cm, regardless of depth or signal strength, showing no observable degradation in accuracy with increasing burial depth. This suggests that the model is not only accurate but also stable across all operating conditions tested. These accuracies translate to actionable guidance for standoff, excavation planning, and rapid triage in reconnaissance workflows where seconds matter.

To further evaluate the influence of background contamination on prediction performance, the mean resolution—with resolution defined as the difference between the predicted and true burial depth—was plotted as a function of the signal-to-total ratio (SNR), as shown in Figure 4. The results demonstrate that the model exhibits minimal bias, with the mean resolution remaining effectively zero across all SNR bins. For SNR values greater than 0.25, the standard deviation of the resolution remains within ~5cm, consistent with the overall RMSE. However, as expected, for spectra dominated by background (SNR = 0.1), prediction uncertainty increases significantly, with standard deviation reaching approximately 12cm. This behavior reflects the difficulty of extracting depth-sensitive features from spectra with minimal signal contribution but also highlights the model's resilience in noisy environments, maintaining usable accuracy even under severe degradation.

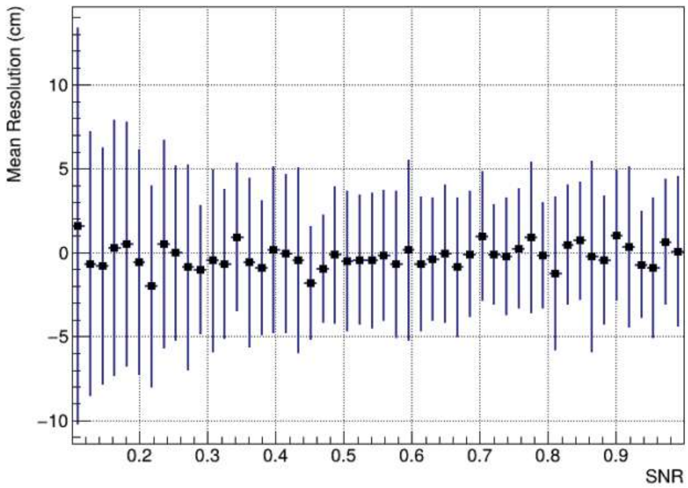

Figure 4 Resolution (defined as predicted depth minus true depth) as a function of signal-to-total ratio (SNR). Error bars indicate the standard deviation of the resolution within each SNR bin. The mean resolution remains near zero across all SNR values, indicating minimal bias. Standard deviations are approximately 5 cm for SNR > 0.25, increasing to ~12 cm only at the lowest SNR (0.1), demonstrating the model's robustness under high-background conditions.

These findings are particularly important when considered in the context of nuclear security applications. Real-world threats such as the attempted trafficking of uranium in Georgia (2025), the interdiction of plutonium and uranium by a Yakuza criminal network (2024), and the repeated losses and thefts of sealed radioactive sources in Mexico and Australia all underscore the need for reliable, field-deployable systems capable of non-intrusively detecting and characterizing buried or shielded sources. In such scenarios, fast and accurate depth estimation is crucial for enabling appropriate response actions, whether by emergency responders, military units, or nuclear forensics teams. The GBDT model presented in this work, by leveraging the full 1024-bin gamma spectrum and capturing subtle features beyond traditional peak-based methods, provides a practical and effective solution for this task.

Conclusions

The gradient boosted decision tree (GBDT) model developed in this work demonstrates a highly accurate and robust method for estimating the burial depth of a ^{137}Cs radioactive source using full-spectrum gamma-ray data. Trained entirely on Geant4-simulated spectra that incorporate realistic energy resolution, statistical fluctuations, and environmental background conditions, the model achieved a coefficient of determination of $R^2 = 0.96$ and a root-mean-square error (RMSE) of 5cm across depths ranging from 5 to 95 cm. As shown in Figure 3, the mean predicted depth closely matches the true burial depth at all levels, with consistent standard deviation errors of approximately 5 cm, regardless of signal-to-total ratio (SNR). Figure 4 further illustrates that the model exhibits minimal prediction bias, with the centered near zero and standard deviations remaining under 5 cm for SNRs greater than 0.25. Only at the lowest SNR level (0.1), representative of highly background-dominated scenarios, does the model's precision decline significantly, with uncertainty rising to approximately 12 cm.

The model's combination of full-spectrum inputs and real-time inference provides immediate value in CBRN reconnaissance: responders can (i) estimate depth to plan digging or shielding strategies, (ii) prioritize sites when multiple hotspots are present, and (iii) communicate quantitative uncertainty (~5 cm at moderate SNR) to inform risk and resource allocation. Integration into handheld or backpack systems is straightforward since only a single NaI(Tl) spectrum is required.

These results have important implications for the field of nuclear security, where non-intrusive detection and localization of radioactive materials are essential for preventing radiological terrorism, illicit trafficking, and public exposure to orphan sources. Real-world incidents, such as the attempted trafficking of uranium in Georgia (2025), the Yakuza-led plutonium smuggling operation thwarted by U.S. authorities (2024), and the repeated thefts and losses of industrial radiological sources in Mexico and Australia, highlight the urgent need for accurate and field-deployable tools to locate hidden or buried radioactive threats. In such scenarios, conventional detection systems may struggle to distinguish signal from background or estimate source depth, particularly when shielding or burial is involved. The machine-learning approach presented here—capable of interpreting the entire gamma spectrum without relying on hand-engineered features—offers a fast, data-driven solution that is both flexible and adaptable to challenging field conditions.

Looking forward, future work will focus on validating the algorithm with experimentally measured spectra, particularly using ^{137}Cs sources buried in real soil under varying moisture and density conditions. Additional extensions will include expanding the model to handle other isotopes of security concern (e.g., ^{60}Co, ^{192}Ir, ^{241}Am), incorporating soil heterogeneity, and investigating sensor fusion with other modalities such as neutron detection. The integration of temporal data from mobile detector platforms and spatial data from sensor arrays also holds promise for improving 3D localization and discrimination of multiple buried sources. Overall, this study contributes to the development of machine learning–based methods for radiological source depth estimation, demonstrating that full-spectrum models can achieve reliable performance under realistic conditions. The results highlight the potential of such approaches to support future tools for non-intrusive assessment in nuclear security and emergency response applications.

References

ABC News. 2023. *Lost Cs-137 Capsule Triggers Radiation Alert in Western Australia.* https://en.wikipedia.org/wiki/Western_Australian_radioactive_capsule_incident.

Agostinelli, S, Allison J., Amako K., Apostolakis J., Araujo H., Arce P., Asai M., et al. 2003. "Geant4—a simulation toolkit." *Nuclear Instruments and Methods A* 506: 250-303. doi:10.1016/S0168-9002(03)01368-8.

Berthold, Michael R., Nicolas Cebron, Fabian Dill, Thomas R. Gabriel, Tobias Kotter, Thorsten Meinl, Peter Ohl, Christoph Sieb, Kilian Thiel, and Bernd Wiswedel. 2007. "KNIME: The Konstanz Information Miner." In *Studies in Classification, Data Analysis, and Knowledge Organization (GfKL 2007)*. Springer.

Friedman, Jerome H. 2001. "Greedy function approximation: A gradient boosting machine." *Ann. Statist.* 29 (5): 1189-1232. doi: https://doi.org/10.1214/aos/1013203451.

IAEA. 2013. *Hijacking of a Cobalt-60 Teletherapy Source Near Mexico City.* https://www.iaea.org/newscenter/news/mexico-says-stolen-radioactive-source-found-field.

Kamuda, Mark, Jifu Zhao, and Kathryn Huff. 2020. "A comparison of machine learning methods for automated gamma-ray spectroscopy." *Nuclear Instruments and Methods in Physics Research Section A* 954: 1-6. https://doi.org/10.1016/j.nima.2018.10.063.

Reuters. 2025. *Georgia Stops Sale of $3 Million Uranium That Could Have Been Used for a Bomb.* https://www.reuters.com/world/europe/georgia-stops-sale-3-million-uranium-that-could-have-been-used-bomb-2025-07-17/.

Shippen, Alan, and Malcolm J. Joyce. 2010. "Profiling the depth of caesium-137 contamination in concrete via a relative linear attenuation model." *Applied Radiation and Isotopes* 68 (4-5): 631-634. https://doi.org/10.1016/j.apradiso.2009.09.046.

U.S. Department of Justice, OPA. 2024. *Justice Department Announces Nuclear Materials Trafficking Charges Against Japanese Yakuza.* https://www.justice.gov/archives/opa/pr/justice-department-announces-nuclear-materials-trafficking-charges-against-japanese-yakuza.

Ukaegbu, Ikechukwu Kevin, and Kelum Gamage. 2018. "A Model for Remote Depth Estimation of Buried Radioactive Wastes Using CdZnTe Detector." *Sensors* 1612. https://doi.org/10.3390/s18051612.

Ukaegbu, S. I., and K. A. A. Gamage. 2018. "A Novel Method for Remote Depth Estimation of Buried Radioactive Contamination." *Sensors* 18 (5): 507. https://doi.org/10.3390/s18020507.

Ukaegbu, S. I., G. O. Aspinall, and K. A. A. Gamage. 2019. "Nonintrusive Depth Estimation of Buried Radioactive Wastes Using Ground Penetrating Radar and a Gamma Ray Detector ." *Remote Sensing* 11: 141. https://doi.org/10.3390/rs11020141.

Funding Information

This research received no external funding.

Conflict of Interest

The author declares no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Data Availability

The data that support the findings of this study are openly available in the Open Science Framework at https://osf.io/36s8x/