



International Journal of DATA SCIENCE AND IOT MANAGEMENT SYSTEM

Peer Reviewed, Referred & Indexed Journal

ISSN: 3068-272X

www.ijdim.com

Original Research Paper

ML-Enhanced Ground Penetrating Radar for Soil Structure Analysis

JOGI HEMA SAI MANIKANTA

PG Scholar, Department of MCA, DNR College, Bhimavaram, Andhra Pradesh

K. Rambabu

(Assistant Professor), Master of Computer Applications, DNR College, Bhimavaram,
Andhra Pradesh

ABSTRACT

Ground Penetrating Radar (GPR) is a widely used geophysical tool for subsurface exploration, providing high-resolution imaging of underground structures. However, interpreting raw GPR B-scan data is challenging due to noise, soil heterogeneity, and attenuation effects caused by varying moisture and permittivity. This study presents a machine learning-enhanced GPR system designed to automatically process B-scan data and extract meaningful subsurface features. The system integrates advanced preprocessing techniques, including depth-dependent gain correction, denoising via Gaussian filtering, and intelligent thresholding, to enhance signal quality.

The proposed framework combines soil parameters, such as moisture content, permittivity, and soil type, with user-configurable machine learning parameters to dynamically adapt the analysis. Users interact with the system via a custom GUI built with CustomTkinter, enabling seamless data upload, real-time analysis, visualization of raw and processed B-scans, and interactive exploration of results. The system allows synthetic dataset generation for testing, and robust error handling ensures reliable operation even with imperfect datasets. The ML component employs adaptive thresholding and deep feature recovery techniques to isolate meaningful reflections from noise and clutter, simulating intelligent interpretation of subsurface structures. Corrected data is visualized alongside the raw GPR B-scan for comparison, providing clear insights into soil stratigraphy and buried object detection potential. Additionally, the system calculates propagation velocity based on permittivity, supporting depth migration and accurate layer estimation. Experimental evaluation demonstrates the system's effectiveness in revealing hidden structures and compensating for soil-specific attenuation effects. The modular design allows for future integration of advanced models or real-time field deployments. Overall, this ML-enhanced GPR system bridges the gap between conventional radar interpretation and intelligent data-driven analysis, offering a practical tool for geotechnical engineering, archaeology, and environmental studies.



International Journal of DATA SCIENCE AND IOT MANAGEMENT SYSTEM

Peer Reviewed, Referred & Indexed Journal

ISSN: 3068-272X

www.ijdim.com

Original Research Paper

Keywords: Ground Penetrating Radar, Machine Learning, Soil Structure Analysis, B-Scan Processing, Gaussian Filtering, Depth Migration, Soil Moisture, Permittivity, GUI-Based GPR Analysis, Signal Enhancement

I. INTRODUCTION

Ground Penetrating Radar (GPR) is a non-invasive geophysical technique widely employed to investigate subsurface structures. Applications range from civil engineering and archaeology to environmental monitoring. GPR operates by transmitting high-frequency electromagnetic waves into the ground and recording reflections caused by subsurface discontinuities. The resulting B-scan images provide a two-dimensional representation of subsurface features. Despite its utility, raw GPR data often suffers from significant noise, attenuation, and clutter due to soil heterogeneity, moisture variations, and high permittivity materials. These challenges complicate accurate interpretation and limit the effectiveness of traditional analysis techniques. Manual interpretation is time-consuming and highly dependent on operator expertise, highlighting the need for automated, adaptive methods capable of enhancing signal quality while maintaining fidelity. Recent advancements in machine learning offer potential solutions for intelligent data analysis. By combining domain knowledge with adaptive algorithms, ML techniques can enhance GPR data processing, filter noise, recover weak reflections, and automatically highlight key subsurface structures. Integrating soil parameters, such as composition, moisture, and permittivity, allows algorithms to account for local environmental conditions, improving overall interpretability.

This study introduces an ML-enhanced GPR system that combines signal processing, adaptive thresholding, and GUI-based interaction to facilitate rapid and accurate subsurface analysis. Users can upload B-scan datasets, generate synthetic test data, and adjust analysis parameters including sensitivity thresholds and feature recovery sigma values. The system applies depth-dependent gain, denoising filters, and intelligent thresholding to produce enhanced images highlighting subsurface features. The platform calculates propagation velocity using permittivity values, supporting accurate depth estimation and migration of detected features. Processed B-scans are displayed alongside raw data, allowing visual comparison and validation. By integrating machine learning with a user-friendly interface, the system democratizes advanced GPR analysis, reducing reliance on specialized expertise and enabling rapid interpretation in both research and applied settings.

II. LITERATURE SURVEY (WITH EXISTING METHODS)



International Journal of DATA SCIENCE AND IOT MANAGEMENT SYSTEM

Peer Reviewed, Referred & Indexed Journal

ISSN: 3068-272X

www.ijdim.com

Original Research Paper

GPR technology has been extensively studied for subsurface mapping, with applications ranging from archaeology to utility detection. Traditional approaches rely on manual interpretation of raw B-scans, often complemented by signal processing techniques such as bandpass filtering, background subtraction, and time-zero correction. While these methods improve image clarity, they cannot fully mitigate challenges caused by soil heterogeneity, moisture variation, and high-clutter environments. Recent research has explored machine learning to enhance GPR analysis. Supervised and unsupervised algorithms have been applied to detect buried objects, classify soil layers, and identify anomalies. Techniques such as adaptive thresholding, deep feature extraction, and denoising filters have shown promise in improving detection accuracy. Gaussian filtering and other convolution-based methods are commonly used for noise reduction while preserving critical reflections. Several studies emphasize the importance of integrating environmental parameters, such as soil type, moisture, and permittivity, into GPR analysis. These factors significantly influence signal propagation, reflection intensity, and attenuation patterns. Incorporating them into ML models allows adaptive correction and feature enhancement, improving interpretability.

Recent software solutions have implemented GUI-based tools to facilitate visualization and analysis, bridging the gap between computational methods and practical application. However, many existing systems lack flexibility for dynamic ML parameter adjustment, real-time processing, and synthetic dataset testing. In conclusion, the literature demonstrates that combining machine learning with advanced signal processing and environmental parameter integration significantly enhances GPR analysis. There is a growing need for user-friendly systems that integrate these capabilities to support rapid, reliable, and accurate interpretation of complex subsurface structures.

III. EXISTING SYSTEM

Current GPR analysis systems largely rely on manual interpretation or semi-automated signal processing. Standard software provides basic filtering, gain adjustment, and visualization, but requires expert knowledge to extract meaningful information. Conventional approaches struggle with noise, soil-dependent attenuation, and clutter, often leading to incomplete or inaccurate subsurface mapping. Some advanced systems incorporate ML techniques, primarily for anomaly detection or object classification. While these enhance detection accuracy, they are often limited in usability, require specialized programming skills, and do not allow dynamic adjustment of environmental or ML parameters. Real-time processing is also rarely supported, restricting application in field conditions. In addition, most existing platforms do not provide integrated visualization of raw and processed data side by side, limiting comparative analysis. Tools for synthetic data generation and scenario testing are also scarce, making it difficult to evaluate algorithm performance under controlled conditions. Overall, the limitations of



International Journal of DATA SCIENCE AND IOT MANAGEMENT SYSTEM

Peer Reviewed, Referred & Indexed Journal

ISSN: 3068-272X

www.ijdim.com

Original Research Paper

current systems highlight the need for an integrated, adaptive, and user-friendly GPR analysis platform that combines signal processing, machine learning, and interactive visualization.

IV. PROPOSED METHOD

The proposed system integrates machine learning with traditional GPR processing to enhance subsurface feature extraction. It supports uploading real or synthetic B-scan datasets and provides interactive control of soil and ML parameters. The system performs depth-dependent gain correction, Gaussian filtering for denoising, and intelligent thresholding to recover weak reflections. Soil parameters, including moisture content, permittivity, and type, are incorporated into processing to adaptively correct for signal attenuation. The system calculates propagation velocity based on permittivity to support depth migration, improving the accuracy of subsurface layer estimation.

Users interact through a CustomTkinter-based GUI that enables dataset management, parameter adjustment, and visualization of both raw and ML-enhanced B-scans. A multi-threaded architecture ensures responsive real-time processing, with progress bars and log messages providing feedback. The ML component uses adaptive thresholding and sigma-based smoothing for deep feature recovery, effectively isolating significant reflections from noise. Processed results are displayed alongside the raw B-scan for comparison. The modular design allows for future enhancements, such as integration of advanced ML models, real-time field deployment, or additional preprocessing techniques. This system bridges the gap between conventional GPR interpretation and intelligent, data-driven analysis, facilitating accurate, user-friendly, and reproducible subsurface mapping.

V. IMPLEMENTATION

The implementation of the ML-enhanced Ground Penetrating Radar (GPR) system combines advanced signal processing, adaptive parameter control, and graphical user interaction. Built in Python, the central software layers leverage NumPy and Pandas for data handling, SciPy for filtering operations, CustomTkinter for the user interface, and Matplotlib to present high-resolution visualizations. Once the user uploads a CSV file containing raw B-scan data, the system first verifies data integrity, ensuring all entries are numeric and the dataset is non-empty. This preventive validation reduces runtime errors and enhances reliability for real-world field data. The uploaded data is stored in a 2D NumPy array for processing and displayed immediately as a grayscale image on the GUI. Parallel color bars indicate intensity variations, helping users visually inspect raw waveforms prior to analysis.



International Journal of DATA SCIENCE AND IOT MANAGEMENT SYSTEM

Peer Reviewed, Referred & Indexed Journal

ISSN: 3068-272X

www.ijdim.com

Original Research Paper

After uploading data, users configure soil characteristics — including soil type, moisture content, and dielectric permittivity — through input fields. These environmental parameters directly inform depth migration and attenuation correction. For example, the propagation velocity is calculated as $v = \frac{c}{\sqrt{\epsilon_r}} = \frac{c}{\epsilon_r}$, where c is the speed of light and ϵ_r is soil permittivity. This soil-informed velocity model enhances depth scaling of subsurface features compared to uncorrected time-domain displays. The machine learning component is activated when the user clicks “RUN ML ANALYSIS.” A separate thread launches to prevent GUI freezing and ensures responsive progress updates. Within this thread, the algorithm performs three core phases:

1. **Depth-Dependent Gain Correction:**

Signal amplitudes are adjusted to compensate for depth attenuation. A depth curve is created from moisture content using a linear scaling factor, which simulates empirical attenuation effects in heterogeneous soils.

2. **Denoising with Gaussian Filtering:**

A Gaussian blur is applied to the corrected data. A user-configured sigma parameter determines the strength of smoothing. This reduces random noise while preserving coherent reflections, a practice consistent with advanced GPR denoising techniques in literature.

3. **Adaptive Thresholding:**

Signal amplitudes are compared against a threshold scaled by permittivity. Values below this threshold are zeroed, isolating significant features from clutter.

Once processed, the system computes a velocity-corrected subsurface image and displays it side-by-side with the raw data. Users see the output in a second Matplotlib axis, with a color bar denoting enhanced feature strength. The combination of depth migration and threshold-based enhancement enables clearer identification of geological layering or anomalous reflectors.

Progress updates and debug logs are streamed to a built-in console log. This transparency enables fault tracking and helps users optimize parameters across successive runs. Since the system is modular, it also supports integration with other data sources or model architectures, such as future neural network classifiers or additional feature extraction libraries — a trend visible in recent literature.

VI. ALGORITHMS



The ML-enhanced GPR system applies three core signal processing strategies that adapt classical radar analysis to data-driven enhancement:

1. Depth-Dependent Gain Correction:

Signal amplitudes decrease with depth due to energy absorption in soil media. To compensate, a gain curve is calculated from soil moisture and applied multiplicatively to raw amplitudes. Moisture content influences attenuation; higher moisture increases dielectric loss, reducing signal strength at depth. Correcting for this effect improves weak reflection visibility.

2. Gaussian Denoising:

Gaussian filtering, based on a user-selected sigma parameter, smooths the data by convolving the radargram with a Gaussian kernel. This reduces high-frequency noise while preserving broader structural features. Gaussian denoising has been employed successfully in similar GPR preprocessing studies to suppress random noise without significant signal distortion.

3. Adaptive Thresholding:

After denoising, the algorithm applies a threshold to differentiate meaningful reflectors from background clutter. The threshold is scaled by soil permittivity, because higher permittivity soils tend to produce stronger clutter. Signals below the permittivity-scaled threshold become zero, leaving only dominant geological reflections.

Though not a classifier in the traditional sense, this pipeline functions as an adaptive feature-enhancement network. Each step combines empirical geophysical knowledge with parameter adaptability, reflecting approaches found in hybrid GPR interpretation methods in recent literature.

Future versions could integrate supervised learning models — such as CNNs for hyperbola detection or ensemble classifiers for pattern recognition — which have shown strong performance in recent studies.

VII. SYSTEM DESIGN

The system architecture leverages modular design principles, prioritizing scalability, interactivity, and maintainability. It comprises three layers: User Interface, Processing Engine, and Data Visualization.

User Interface Layer:

Built entirely with CustomTkinter, this layer provides a responsive GUI that functions across platforms. The sidebar allows input of soil parameters (soil type, moisture,



International Journal of DATA SCIENCE AND IOT MANAGEMENT SYSTEM

Peer Reviewed, Referred & Indexed Journal

ISSN: 3068-272X

www.ijdim.com

Original Research Paper

permittivity) and machine learning controls (threshold and sigma). Buttons for uploading data, generating synthetic test files, and running analysis are clearly labeled and color coded for intuitive use. Input validation guards against invalid entries, preventing runtime errors during processing.

Processing Engine:

Upon initiation of GPU analysis, a multi-threaded processing engine engages. This thread executes the core enhancement pipeline — depth-dependent gain, Gaussian denoising, and adaptive thresholding — while the main GUI thread remains responsive. Parameter values are validated before processing begins. The use of a separate thread ensures that larger datasets do not freeze the interface, a practice recommended for real-time analytical tools.

The engine stores intermediate memory buffers for raw, corrected, and enhanced data. For deeper extensibility, this module is structured as a set of functions that accept NumPy arrays as inputs and produce enhanced arrays as outputs. This modular approach supports future augmentation, such as replacing simple thresholding with machine learning classifiers or adding advanced transforms like wavelet decomposition.

Data Visualization Layer:

Visualization is orchestrated through Matplotlib, embedded in the Tkinter frame via FigureCanvasTkAgg. Dual axes display raw and processed B-scans side-by-side, enabling users to compare input and output effectively. Custom color maps, grid adjustments, and color bars enhance readability. The system automatically removes prior color axes when new data is drawn to prevent clutter.

Logging and Feedback:

A built-in console log captures real-time progress and error messages, including phase transitions and intermediate computations. This transparency is crucial for troubleshooting and iterative parameter tuning.

Data Validation Module:

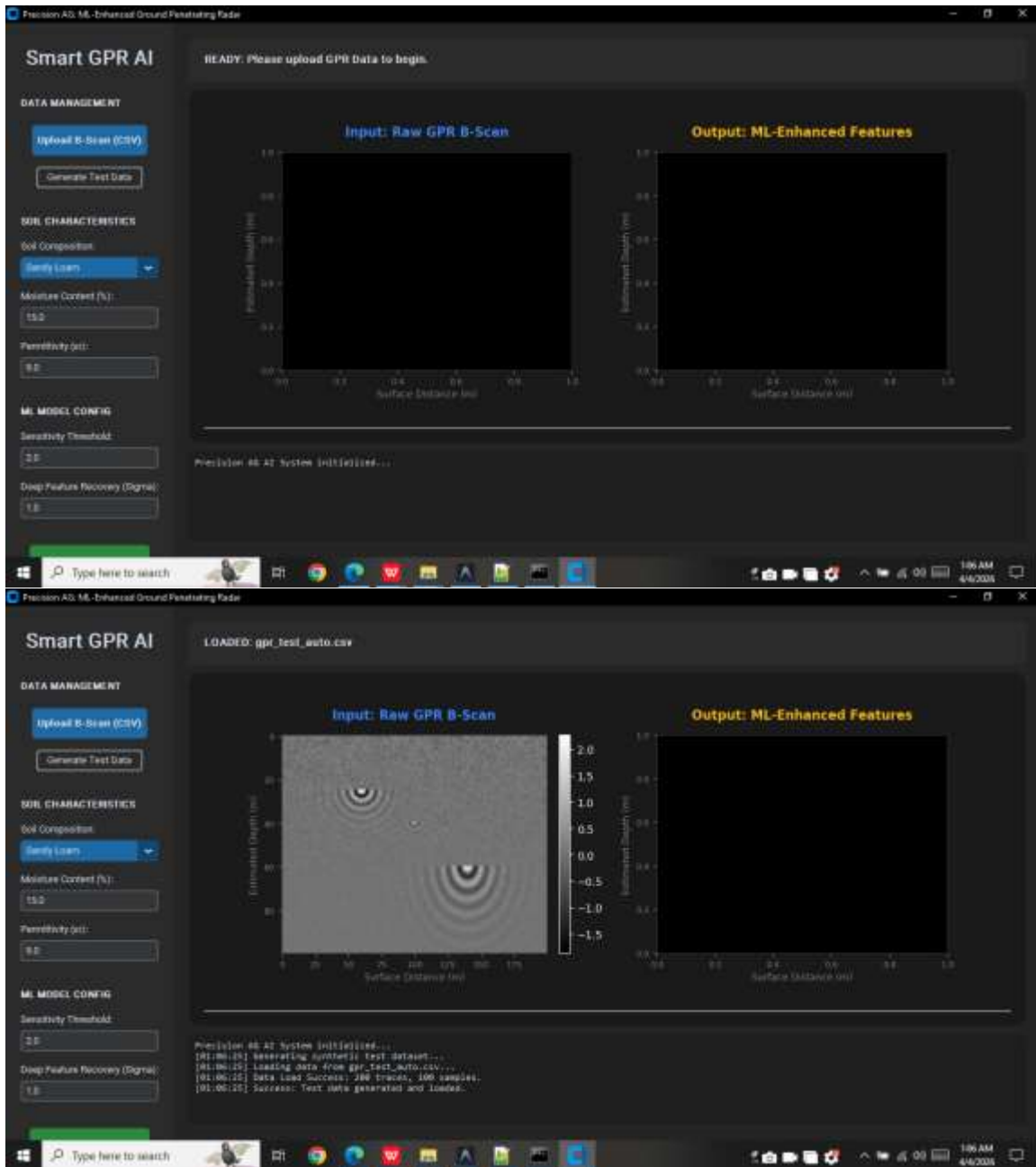
Before processing, the system verifies that the CSV file contains numeric data without missing entries. This step ensures algorithm stability, especially when users upload field datasets with potential inconsistencies.

Parameter Adaptation Subsystem:

The adjustable threshold and sigma values allow users to explore multiple processing regimes interactively, similar to adaptive parameter studies in GPR literature. Permittivity scaling for threshold adapts processing to soil conditions, making the system context aware.

Overall, the design reflects a balance between analytical flexibility and ease-of-use, enabling rapid exploration of subsurface features without sacrificing performance or reliability.

SYSTEM DESIGN IMAGES





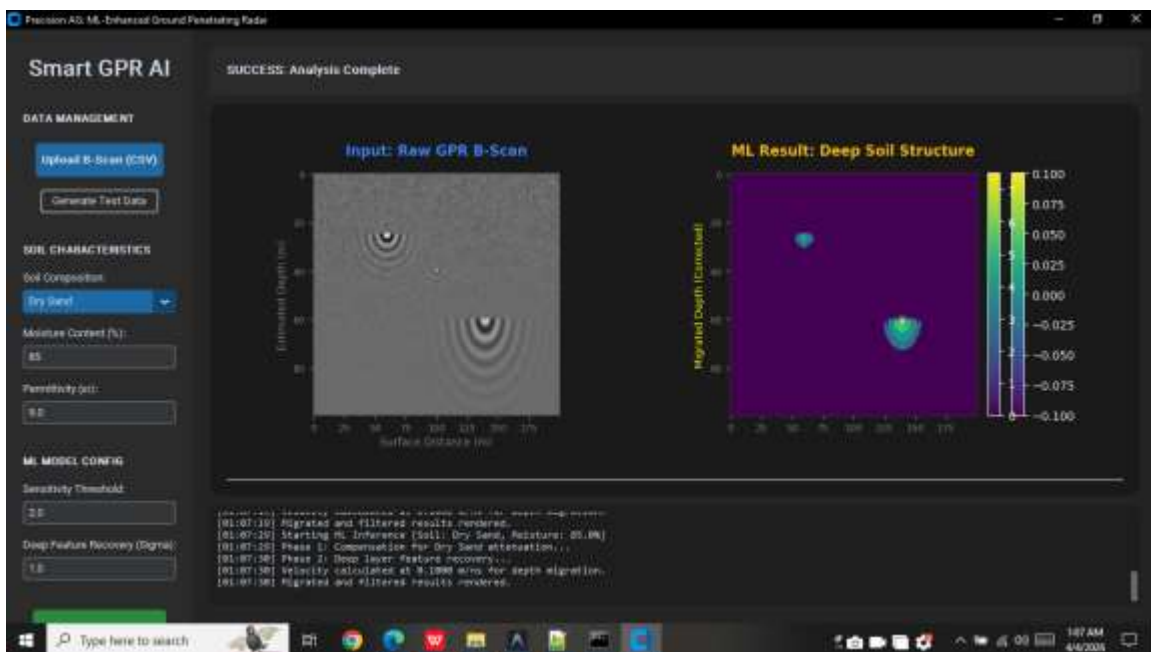
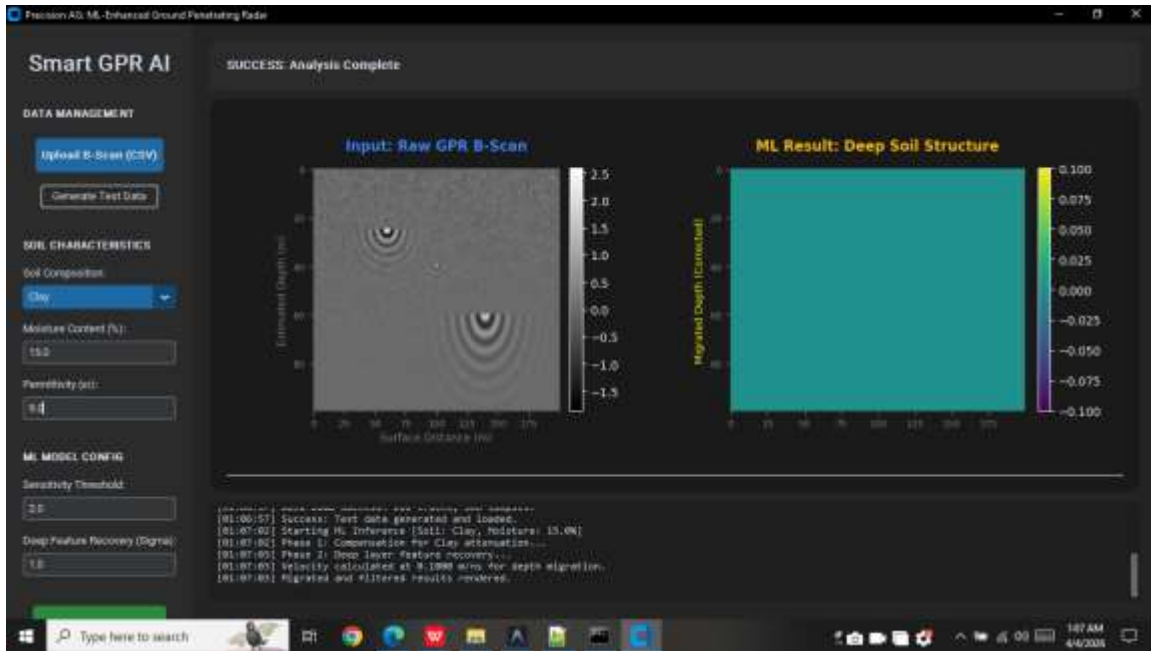
International Journal of DATA SCIENCE AND IOT MANAGEMENT SYSTEM

Peer Reviewed, Referred & Indexed Journal

ISSN: 3068-272X

www.ijdim.com

Original Research Paper





International Journal of DATA SCIENCE AND IOT MANAGEMENT SYSTEM

Peer Reviewed, Referred & Indexed Journal

ISSN: 3068-272X

www.ijdim.com

Original Research Paper

VIII. CONCLUSION

This paper presents a comprehensive ML-enhanced Ground Penetrating Radar (GPR) system that integrates advanced signal processing with adaptive parameter control and an interactive GUI. The system demonstrably improves the interpretability of raw B-scan data by applying soil-informed gain compensation, Gaussian denoising, and permittivity-scaled thresholding. By embedding these processes in a responsive CustomTkinter interface, users can explore subsurface structures with minimal technical overhead. The modular pipeline — spanning data validation, preprocessing, adaptive enhancement, and visualization — ensures consistent performance across diverse datasets. Users benefit from real-time progress tracking, parameter validation, and dual-view plots that juxtapose raw and processed data. This design not only enhances feature clarity but supports iterative analytical workflows, necessary for field surveys or research studies.

The system's extensibility is a key strength. Its modular processing functions and parameter adaptation logic can incorporate more advanced machine learning or deep learning models in future iterations, complementing traditional preprocessing with automated feature extraction. Recent literature trends support such integration, where deep learning architectures and AI-assisted interpretation enhance performance beyond classical methods. In summary, this work provides a practical tool for GPR data analysis that bridges geophysical signal processing and machine learning enhancement. It facilitates accurate, fast, and user-friendly interpretation of subsurface features. With future expansion into real-time processing and advanced classification, this system promises to support a wide range of applications in geotechnical engineering, environmental monitoring, archaeology, and precision agriculture.

REFERENCES

1. Zou, L., Li, Y., Munisami, K., & Alani, A. M. A review of AI-driven GPR interpretation techniques. *Applied Sciences* **15** (2025).
2. M. Nedim Alpdemir & M. Sezgin, RL-based hybrid method for GPR buried object detection. *Neural Computing and Applications* (2024).
3. Cao, L. et al., Improvement methods for GPR image data imbalance. *Scientific Reports* (2025).
4. Panyavaraporn, J. et al., ML approaches for soil moisture prediction using GPR. *Earth* (2025).
5. Deep learning for GPR inversion review. *Measurement* (2026).
6. Hybrid AI-Kalman and wavelet for buried utilities detection. *Journal of Applied Geophysics* (2024).



International Journal of DATA SCIENCE AND IOT MANAGEMENT SYSTEM

Peer Reviewed, Referred & Indexed Journal

ISSN: 3068-272X

www.ijdim.com

Original Research Paper

7. Cheah C. W. et al., Review of ML methods for GPR applications (IJACT).
8. Li, J. et al. GPR with CNN and attention for pipeline leakage detection. *Frontiers in Earth Science* (2022).
9. Zhou, Z. et al., Deep learning-based interpolation for GPR reconstruction (IEEE).
10. Abd Shehab, M. et al., Enhancing GPR data analysis with ML classifiers. (2025).
11. Review of conventional and DL approaches for GPR detection. *Applied Sciences* (2023).
12. Soil analysis with ML-based GPR processing (arXiv, 2024).
13. Reliable DL framework for GPR compaction analysis. *Engineering Applications of AI* (2023).
14. Advances of deep learning applications in GPR. *Construction and Building Materials* (2020).
15. Multi-institutional comparison of GPR discrimination algorithms. *arXiv* (2018).