

# Size estimation of underground targets from GPR frequency spectra: A deep learning approach

Nairit Barkataki<sup>a,\*</sup>, Banty Tiru<sup>b</sup> and Utpal Sarma<sup>a</sup>

<sup>a</sup>Department of Instrumentation & USIC, Gauhati University, Guwahati, India

<sup>b</sup>Department of Physics, Gauhati University, Guwahati, India

---

## ARTICLE INFO

### Keywords:

object size estimation  
object detection  
FFT  
deep learning  
ground penetrating radar

## ABSTRACT

GPR (Ground Penetrating Radar) is a robust and effective device for identifying underground artefacts. Construction companies and civil engineers should be aware of the sizes of rebars and pipelines before and during construction work for various reasons. Most research efforts have typically concentrated on GPR signal analysis in the time domain, however recent studies have increasingly focused on analysis in the frequency domain. This paper proposes an artificial neural network (ANN) model for estimating the diameter of underground rods (solid) and pipes (hollow). GPR data captured in the time domain domain is transformed to the frequency domain using FFT, after which feature extraction is performed using ANN. An FPGA-based prototype GPR system is used to collect GPR A-Scan data for a variety of targets made of aluminium, stainless steel, rebar, and PVC. A mean absolute percentage error of 1.89% is achieved using the proposed model. The experimental results confirm the effectiveness of the proposed approach in extracting size-related information from GPR data.

---

## 1. Introduction

Ground Penetrating Radar (GPR) is a versatile and reliable tool that is used for detecting buried objects. Its major strengths include its low cost compared to traditional methods, and non-destructive nature. There are numerous applications of GPR, including estimation of soil, snow and ice properties (Jol, 2008; Wu et al., 2019; Liu et al., 2021), non-destructive inspection of archaeological sites (Miccinesi et al., 2021), examining building condition (Daniels, 2004), inspecting roads and tunnels (Persico, 2014), detecting landmines and other buried explosive devices (Daniels, 2004), borehole inspection (Solla et al., 2021), study agriculture properties of soil (Akinsunmade et al., 2019), locating weak zones within embankments (Tomecka-Suchoń, 2019) etc.

A GPR survey often takes much lesser time than alternative technologies. When compared to other geophysical tools such as seismic, transient electromagnetic, electrical, and magnetic techniques, GPR gives data that has higher resolution and better accuracy (Benedetto and Pajewski, 2015).

Planning and conducting an effective GPR survey requires a lot of experience. The interpretation of GPR data is usually not intuitive and special skills are required to convert the measurements into clear images that can help in making critical decisions in various fields like civil engineering, archaeology, military etc.

In case of GPR-based concrete, bridge and road investigations, the most common applications appear to be the detection of pipes (Ayala-Cabrera and Izquierdo, 2021; He et al., 2021) and reinforcement bars (Wang et al., 2020), identifying defects in roads and bridges (Rasol et al., 2022), and to determine the thickness of pavements (Zhao and Al-Qadi, 2017).

One area where significant improvement has been realised is in the estimation of the size of buried targets. Several researchers have published their work on estimating rebar diameter (Mechbal and Khamlichi, 2017; Pasculli et al., 2018; Sun et al., 2021) However, the signal processing strategies proposed by them have proven to be complex and time intensive.

When the diameter of a pipe is smaller than the wavelength of the signal, estimating its diameter becomes challenging. In such cases, the arrival times of the echo signal cannot be co-related to the the diameter of the pipe (Jazayeri et al., 2019). This is typical in underground pipe detection, where the pipe diameter may be 50 mm and the wavelength

---

\*Corresponding author

✉ [nairitb@gauhati.ac.in](mailto:nairitb@gauhati.ac.in) (N. Barkataki)

ORCID(s): 0000-0002-7988-072X (N. Barkataki); 0000-0003-1761-0983 (B. Tiru); 0000-0001-7888-1088 (U. Sarma)

of the transmitted pulse is 75 cm (for 400MHz centre frequency antenna). A variety of solutions have been proposed to overcome this drawback.

GPR does not directly measure the diameter of rebars, cables or pipes (Rathod et al., 2019). When used with targets having larger diameters, GPR can produce stronger radar wave reflections. However, depending on the wavelength of the signal, smaller objects may produce very low intensity reflections that may seem like a dot in the final B-Scan image. Rathod et al. (2019) presented a study where they tried to estimate rebar spacing and diameters using GPR. They were able to estimate rebar spacing with errors ranging from 2-11%. However, the percentage errors in determining rebar diameter was beyond 100% in some cases.

In other studies, researchers have used a combination of GPR and an EM-based system to detect and estimate the size of rebars. Zhou et al. (2018a) developed a GPR-EMI dual sensor based system that simultaneously estimated rebar diameter and cover thickness. Barrile and Pucinotti (2005) combined GPR with a different commercial EM-based system to detect rebars. They reported that they could estimate rebar diameter with a 12% error. This method, however, was time-consuming as data collection using the handheld concrete pachometer was slow. A linear relationship was found between the rebar diameter and maximum GPR amplitude by Hasan and Yazdani (2016a). However, it is not possible to quantitatively relate the peak amplitude of the echo signal with the rebar diameter without knowing the amplitude of the source wavelet and concrete conditions.

The quality of GPR data is usually degraded by background clutter (Kumar et al., 2021), noisy environments like roots and bricks (Zhang et al., 2021), heterogeneous soils and mutual interactions of the waves (Lei et al., 2019).

Researchers have provided a variety of ways for analysing GPR data in recent years. Histograms of Oriented Gradients (HOG) based feature extraction method has been used effectively to identify landmines (Torriero et al., 2013). Fourier Transform and Principal Component Analysis have also been used to identify landmines (Ko et al., 2012). Diameter estimation of underground pipelines and wires by fitting hyperbolas and localising hyperbola coordinates are widely used methods we well (Wahab et al., 2013; Mertens et al., 2015; Zhou et al., 2018b; Jazayeri et al., 2019; Lei et al., 2019; Liu et al., 2020).

Other techniques begin with machine learning algorithms that extract hyperbolic features, followed by a fitting algorithm. Deep-learning-based interpretation algorithms have been the subject of extensive study in recent years (Gamba and Lossani, 2000; Shaw et al., 2005; Ahmed et al., 2021).

Among the earliest proponents of using neural networks in GPR data interpretation, Costamagna et al. (1998) used neural networks to search for object signatures from GPR data. They demonstrated that pipe signatures extracted using neural networks had the same degree of precision as human operators. To determine the position of buried items in GPR pictures, Singh and Nene (2013) used a curve fitting technique along with neural network. The buried objects were localised with curve fitting technique. Their proposed approach could locate buried objects with a total reported accuracy of 91%. The adoption of the deep learning neural network idea for GPR underground utilities has demonstrated promising results (Luo et al., 2020; Tong et al., 2020; Barkataki et al., 2022).

Lei et al. (2019) utilised a R-CNN to isolate potential hyperbolic regions from gray GPR B-scan images. The detected rectangular region was then converted to a binary image and hyperbolic signatures were identified using a double cluster seeking estimate algorithm. Their proposed scheme could automatically extract information related to buried objects with a precision of 96% on real GPR data. Lei et al. (2020) proposed a scheme to estimate the diameters of buried objects. Their method involved an adaptive target region detection (ATRD) algorithm to isolate the hyperbolic signatures in B-Scans, followed by a CNN model for feature extraction and estimation of diameter. Their framework achieved an accuracy of 92.5% on field datasets.

The majority of research efforts have traditionally focused on the analysis of GPR signals within the time domain, although recent studies have increasingly focused on analysis in the frequency domain. Laurens et al. (2005) analysed the relationship between the spectral content of signals reflected by concretes of various degrees of saturation, using a 1.5 GHz antenna system. They found a difference in the central frequency of the reflected signals when the spectrum for dry (1.4 GHz) and saturated (1.18 GHz) concrete were compared.

Studies by Rodés et al. (2015) have shown that there is a distinct correlation between the spectrum of the recorded traces and the subsurface structures. Radar signals are scattered when targets are present in the medium. The amplitude of the wave,  $A(r)$  at a certain distance  $r$  from the source, can be represented as a function of the wave amplitude at the source and the losses caused by those attenuating effects:

$$A(r) = A_0 \frac{1}{r} e^{-(\alpha+\mu)r} \quad (1)$$

$A_0$  being the amplitude of reference at the source,  $r$  the distance,  $\alpha$  the coefficient of attenuation due to the absorption and  $\mu$  the coefficient of scattering.

The coefficient of scattering depends on the number of scatterers per unit of volume,  $n$  and on the cross section of the scatterers,  $C_s$ .

$$\mu = \frac{nC_s}{2} \quad (2)$$

Moreover, the effect of scattering is frequency dependent, because the cross section depends on the frequency of the incident field (Salinas Naval et al., 2018).

Estimating the size of the target object is one area where considerable progress has been made. Several research groups have established experimental relationships between the amplitude and frequency content of the hyperbolic returns and the diameters of embedded bars, as well as additional properties of the rebar and concrete (Kalogeropoulos et al., 2011, 2012; Lai et al., 2013; Hasan and Yazdani, 2016b).

Bi et al. (2018) analysed the spectral characteristics of target GPR signals and proposed a method to suppress random noise from GPR data. They used a singular value decomposition (SVD) method to filter out unwanted data and reconstruct the GPR data. They found that the proposed approach showed improved results in suppressing noise and clutter, compared with the conventional time-domain method.

According to the study presented by Che et al. (2021), it was easier to detect the depth and position of buried objects when analysing the GPR data in frequency domain. Park et al. (2021) applied frequency-wavenumber migration (f-k migration) using the Fast Fourier Transform (FFT) to the GPR data and used YOLO-v3 algorithm to detect rebars and estimate their diameters. Three rebars were used having diameters 10 mm, 20 mm and 30 mm. Their proposed model could estimate rebar diameters with an average accuracy of 98% and was found to be more accurate than other conventional methods. GPR data was also analysed in the frequency domain by Kang et al. (2022) to detect and estimate the size of a cavity beneath a concrete plate.

However, there is still scope of improvement in the accuracy, speed, and other parameters of the models proposed by researchers. This study presents an novel artificial neural network (ANN) based model to estimate the diameter of buried rods (solid) and pipes (hollow). GPR data collected in the time domain domain is converted to frequency domain using FFT, after which the ANN-based estimator is used. Data is collected using a Field Programmable Gate Array (FPGA) based prototype GPR system.

## 2. Materials and Methods

### 2.1. Hardware Platform for the GPR System

FPGA-based embedded systems provide a rapid prototyping platform, given the flexibility of these devices for intricate system designs, as well as low static and dynamic power consumption, high performance and a re-programmable medium to implement a variety of applications.

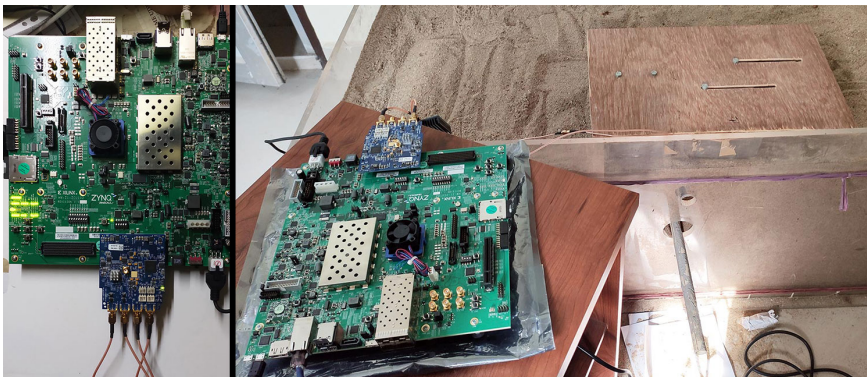
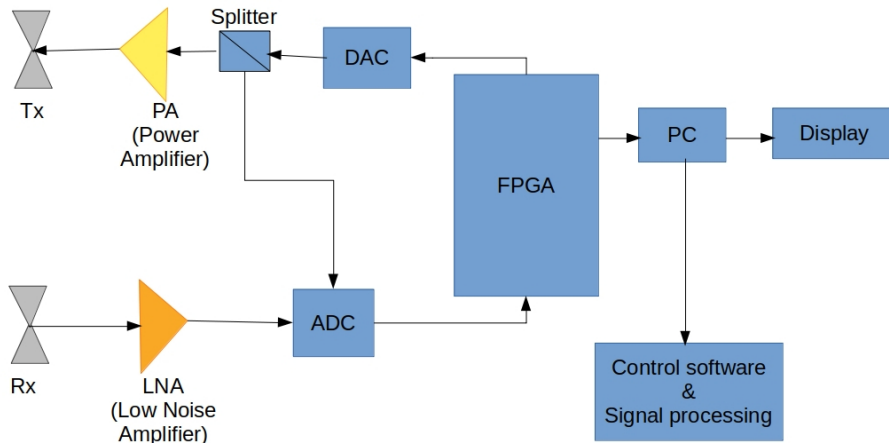


Figure 1: Signal processing platform with test setup

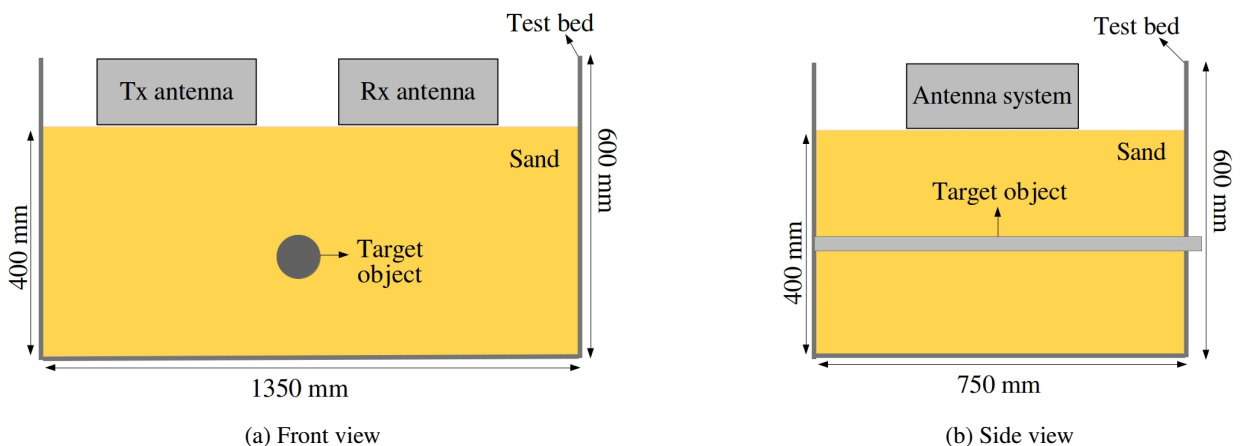


**Figure 2:** Block diagram of the system

A hardware platform for the prototype GPR system has been developed which uses a Xilinx UltraScale+ MPSoC FPGA board as the base. A 14-bit, 1.0 GSPS ADC and a 16-bit, 2.8 GSPS DAC are interfaced to the FPGA board for both signal generation (transmission) and digitisation of the received signals. A ricker, gaussian or any custom waveform can be generated using the DAC. This signal (waveform) is then amplified using a power amplifier before feeding it to a transmitting antenna. Two custom designed 400 MHz coaxial-fed bow-tie antennas are used as transmitting and receiving antennas (Barkataki et al., 2021). The received signal is first amplified using a low noise amplifier (LNA) before feeding the signal to the ADC. The digitised signal is then processed by the FPGA and various GPR signal processing algorithms are applied before a final image (A-Scan / B-Scan) can be formed (see Figure 1). A block diagram of the hardware system is shown in Figure 2.

### 2.1.1. Experimental Setup

A test bed of dimension  $1350\text{ mm} \times 750\text{ mm} \times 600\text{ mm}$  ( $X \times Y \times Z$ ) is constructed for collecting the experimental data. It is filled with sand till the height of 400 mm from the bottom of the test bed. Holes are made on one side, through which the target objects are inserted. The antennas are placed on the surface of the sand to collect the data. The collected data are sampled by the on-board ADC at 1 GHz which gives a single sample (data point) in 1 ns. A



**Figure 3:** Diagram of the experimental setup

diagram of the experimental setup is shown in Figure 3. A picture of the constructed test bed used for collecting experimental data is shown in Figure 4.

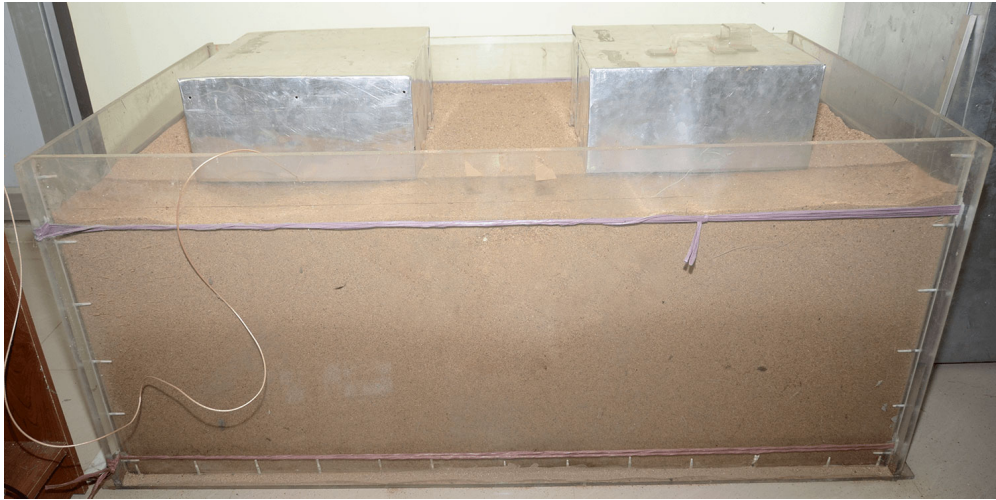
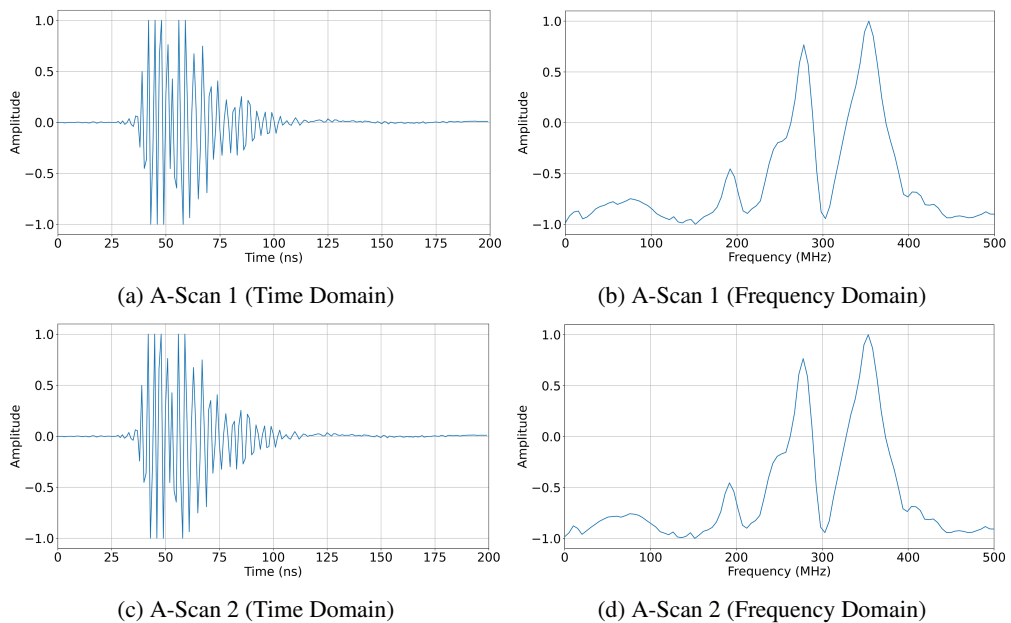
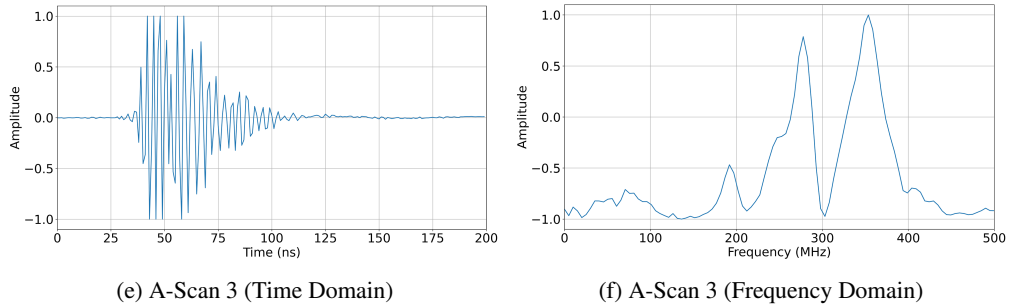


Figure 4: The test bed used for collecting experimental data

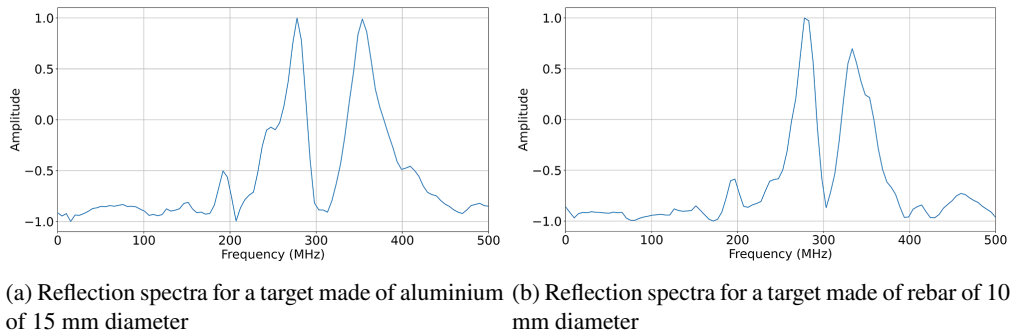
### 2.1.2. Validation of Hardware Platform

A verification of the hardware platform for the GPR system is necessary to ensure that it can provide reliable results before starting the experiments for this study. Initially, multiple GPR A-Scans without any target object are collected for 199 ns, which gives 200 data points from 0<sup>th</sup> second to 199<sup>th</sup> second for a single A-Scan. As can be seen in Figure 5, the spectra of the received wavelets are similar. The signal in the time domain is also similar. Later, GPR A-Scans are also collected for different buried targets. Analysis of the received wavelets indicates a marked difference in the spectrum of the signal for buried targets, as shown in Figure 6. Having verified the system's ability to generate reliable data, experimental data collection is started.

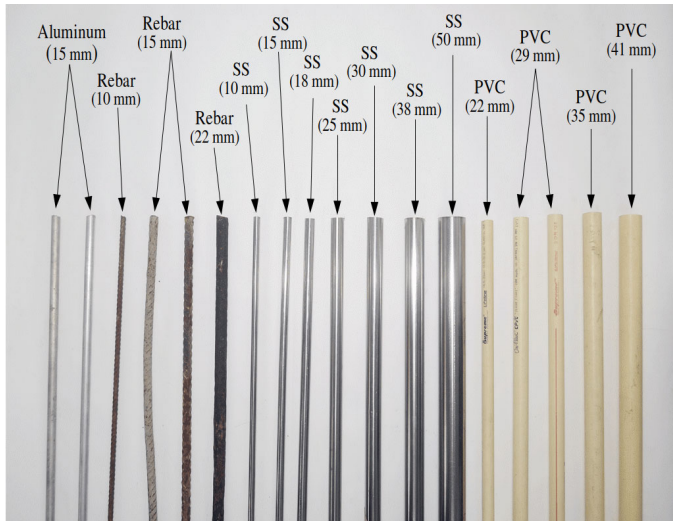




**Figure 5:** Received wavelets for no buried targets



**Figure 6:** Spectrum of received wavelets for buried targets



**Figure 7**  
The different bars and pipes used for the study

**Table 1**  
Targets used in this study

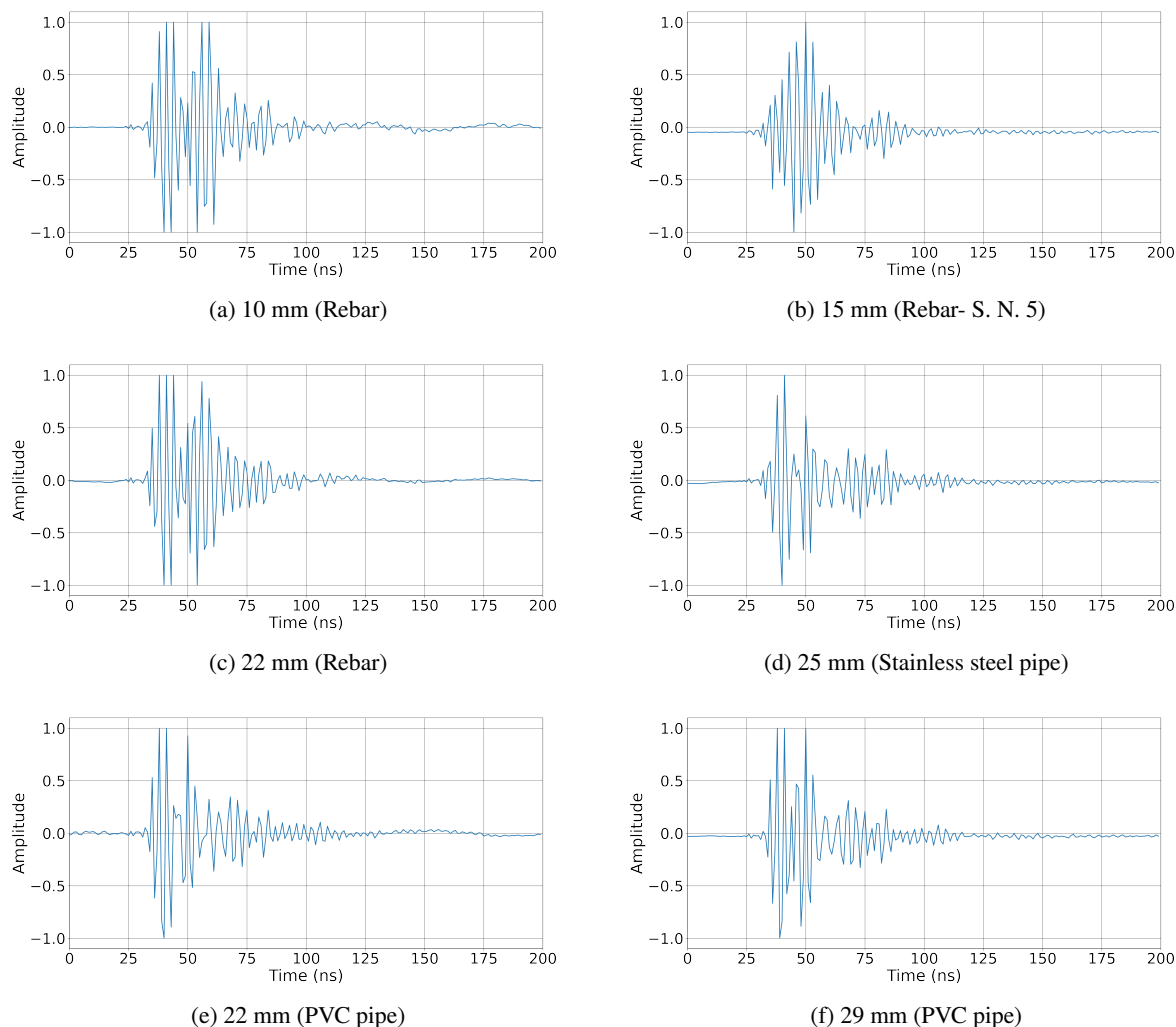
Sl. No.	Target	Outer Diameter (mm)
1	Aluminium bar	15
2	Aluminium bar	15
3	Reinforcement bar	10
4	Reinforcement bar	15
5	Reinforcement bar	15
6	Reinforcement bar	22
7	Stainless steel pipe	10
8	Stainless steel pipe	15
9	Stainless steel pipe	18
10	Stainless steel pipe	25
11	Stainless steel pipe	30
12	Stainless steel pipe	38
13	Stainless steel pipe	50
14	PVC Pipe	22
15	PVC Pipe	29
16	PVC Pipe	29
17	PVC Pipe	35
18	PVC Pipe	41

## 2.2. Materials

The buried objects are made of aluminium, stainless steel (SS), reinforcement bar (rebar), and polyvinyl chloride (PVC) with diameters as mentioned in Table 1 and shown in Figure 7. The depth of the buried objects are varied from 60 mm to 140 mm from the soil surface.

### 2.2.1. Case 1: Data in Time Domain

Experimental data is collected using the prototype GPR system described earlier. A total of 1713 A-Scans are collected having 200 data points each. Some collected A-Scans for target of different diameters are shown in Figure 8.

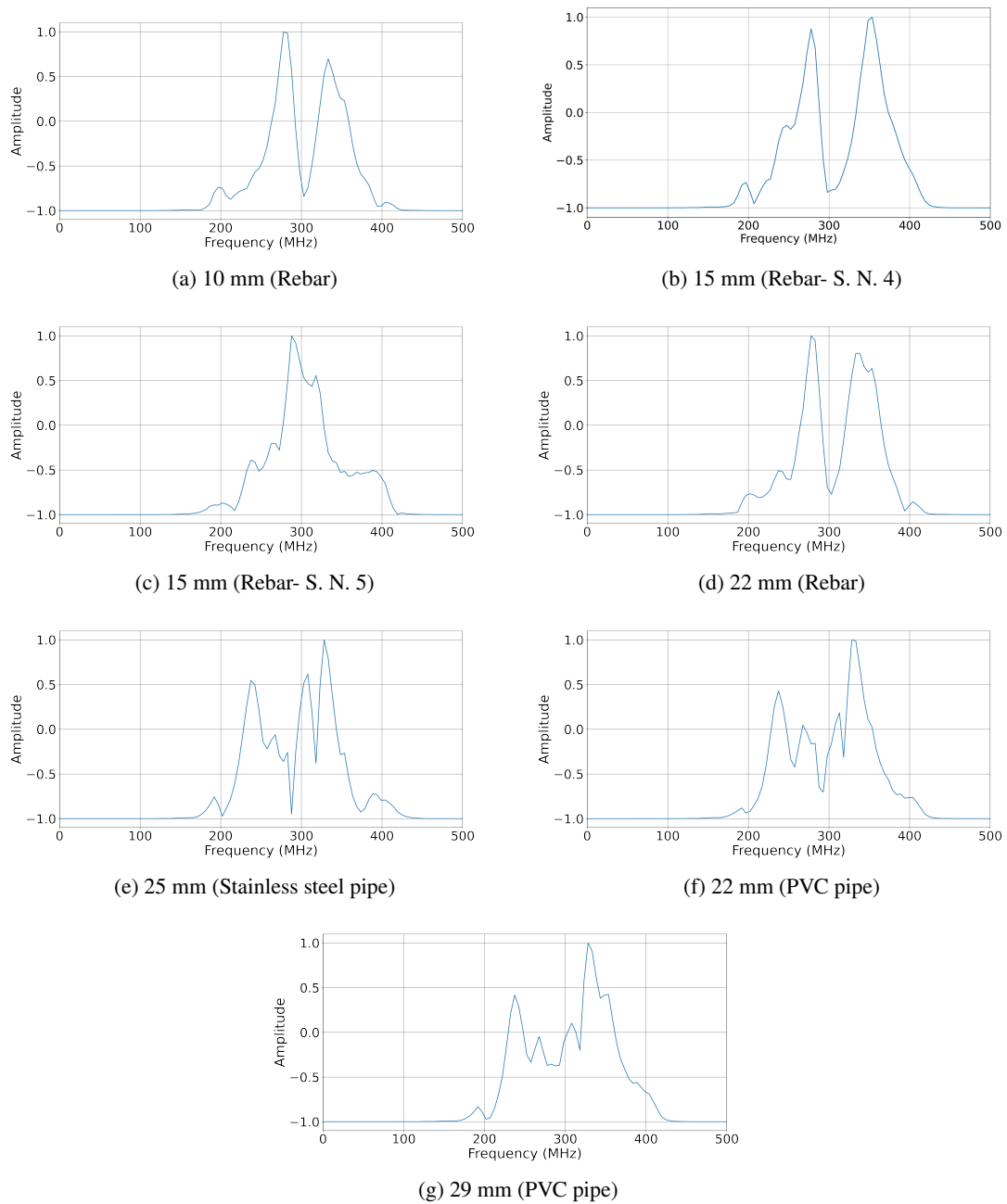


**Figure 8:** Received wavelets for various targets having different diameters

### 2.2.2. Case 2: Data in Frequency Spectrum

After analysing the frequency spectrum of the collected A-Scans, it is seen that the dominant spectrum for all cases ranges from 200-400 MHz. A band-pass filter of 200-400 MHz is applied to each GPR A-Scan, before further processing. Figure 9 shows the spectrum of A-Scans after application of the band-pass filter.

From the above figures, it is clear that there is a co-relation between the spectrum and the size of the buried objects, as was observed by other researchers (Santos et al., 2014; Rodés et al., 2015; Salinas Naval et al., 2018; Pajewski et al., 2019). However, it is seen that the spectra for the rebars of 15 mm (S. N. 4 and 5 in Table 1) are slightly different. This may be due to the difference in shape between the two. Very little or no difference in the spectra is seen for targets made of aluminium (S. N. 1 and 2) and PVC (S. N. 15 and 16) of different qualities. Data is collected using multiple targets of the same diameter value for data augmentation. This is to ensure that small changes in the object's orientation do not affect the trained neural network model. Two separate datasets are created from the GPR data containing information in time domain (TD-dataset) and frequency domain (FD-dataset).



**Figure 9:** Frequency spectrum of received wavelets for various targets

### 2.3. Proposed ANN architecture

An ANN model is used in this study, where feature extraction occurs hierarchically. It is composed of several layers, each containing numerous neurons. Neurons in one layer are linked to every other neuron in the next layer. ANN comprises mainly forward and backward propagation. In forward propagation, each neuron is assigned weights and biases, as well as an activation function. The error between true and predicted values is calculated using a loss function. To determine the optimal hyper-parameters, backward propagation uses an optimisation function to minimise the loss in each successive iteration during training of the model.



**Table 2**  
Validation Losses for different numbers of hidden layers

Sl. No.	No of Hidden Layers	Validation Loss (MAPE)	
		Time Domain	Frequency Domain
1	1 Layer	15.63	8.17
2	2 Layers	12.60	3.56
3	3 Layers	9.84	2.61
4	4 Layers	8.78	2.14
5	5 Layers	8.46	2.42
6	6 Layers	7.34	1.89
7	7 Layers	6.72	2.08
8	8 Layers	6.02	2.18
9	9 Layers	5.67	2.28
10	10 Layers	4.70	2.00

Out of a total of 1713 A-Scans, 1456 (85%) are used for training and 257 (15%) are used for validation of the model. The dataset is normalised before training the model.

### 2.3.1. Loss Function

Mean absolute percentage error (MAPE) is used as a loss function during training of the model. It is a measure of the absolute difference between the true and predicted values, and is expressed as a percentage. It is defined as the mean of absolute relative errors:

$$MAPE = \frac{1}{N} \sum_{t=1}^N \left| \frac{P_t - A_t}{A_t} \right| \quad (3)$$

where N is the no. of samples, ( $P_t$ ) is the predicted value and ( $A_t$ ) is the true value.

### 2.3.2. Evaluation Metrics

MAPE is utilised as a metric to evaluate the training and validation of the ANN model. MAPE is a measure of a regression model's performance that can be expressed in percentages, making it easily understood by anyone.

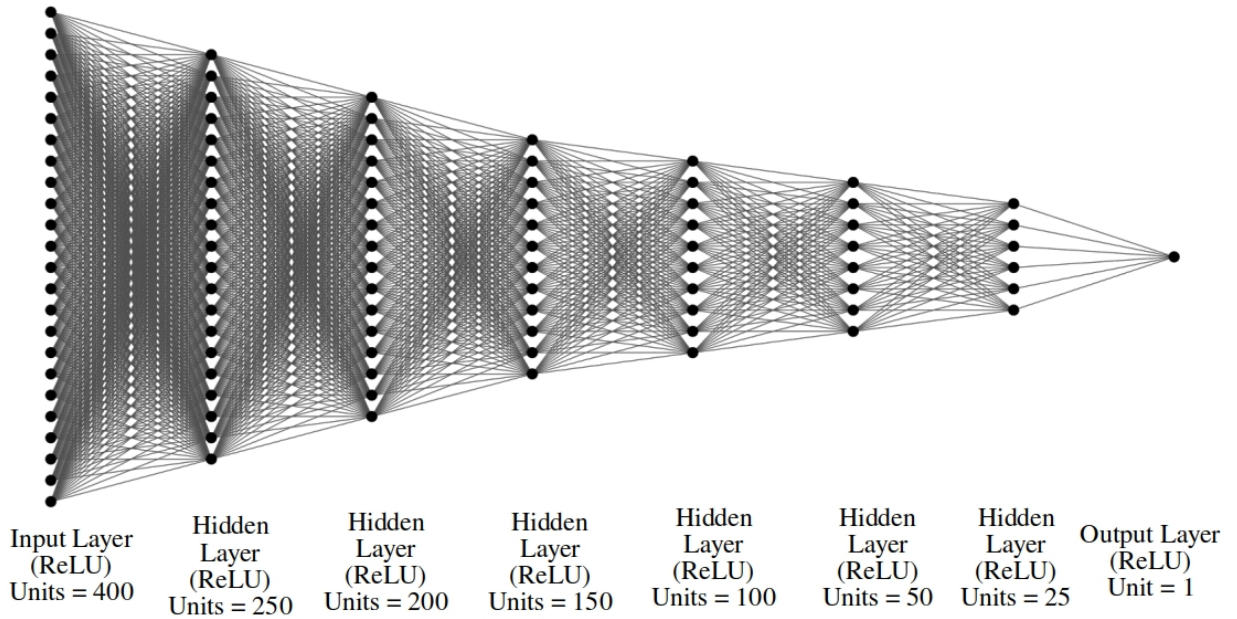
### 2.3.3. Final Architecture

Initially, 200 neurons are used in the input layers and multiple hidden layers are added which have half the number of neurons than the previous layer. Performance is evaluated for different number of hidden layers. Same procedure is repeated for both the datasets containing GPR data in time domain and frequency domain. Table 2 shows the validation losses (MAPE) corresponding to different numbers of hidden layers for both datasets. From the table, it is clear that training the model on the dataset containing GPR spectrum information outperforms that one containing time domain data.

## 3. Results and Discussion

Multiple training runs show that the ANN model with 6 hidden layers and the input layer containing 400 neurons has the best validation loss of 1.89% using spectrum information from the collected GPR data. The proposed model consists of 7 layers including the input layer. ReLU activation function is used in all the layers. To estimate the size of the target, the final output layer with 1 unit uses ReLU as the activation function. Adam optimisation algorithm is used with an initial learning rate of 0.00001. All the models are trained for 2000 epochs and NVIDIA RTX3090 GPU is used during training and validation. The architecture of the best ANN model is shown in Figure 10.

257 A-Scans are used for evaluating the performance of the model. The best and the worst predicted values for different diameters of buried objects are shown in Table 3a. Different metrics like mean squared error (MSE), mean absolute error (MAE), mean absolute percentage error (MAPE), and mean squared logarithmic error (MSLE) are used to analyse the model's performance, as given in Table 3b. It is seen that the model has a MSE of 1.017, MAE of 0.457 and MSLE of 0.0017.



**Figure 10:** Architecture of the proposed ANN

**Table 3**

Performance of the ANN model on the validation set

(a) Predicted values					(b) Performance metrics	
Sl. No.	True values (mm)	Predicted values (mm)		Median (mm)	Metrics	Value
		Worst	Best			
1	10.0	10.84	9.99	9.99	MSE	1.017
2	15.0	14.17	15.00	14.97	MAE	0.457
3	18.0	18.70	17.98	18.11	MAPE	1.887
4	22.0	14.77	22.05	22.10	MSLE	0.0017
5	25.0	26.81	25.05	25.05		
6	29.0	24.40	29.00	28.83		
7	30.0	31.06	30.02	29.94		
8	35.0	36.86	35.03	34.88		
9	41.0	41.96	41.04	41.04		
10	38.0	39.43	37.95	37.95		
11	50.0	46.82	50.04	49.86		

### 3.1. Testing on unseen data

Repeatability is a major concern for any algorithm or instrument. Hence, it is essential to test a ANN model on new and unseen data to be able to properly evaluate its performance.

The experimental setup is used to collect 191 additional A-Scans for the final performance evaluation of the proposed model. When taking measurements, targets of different diameters made of aluminium, rebar, stainless steel, and PVC are picked at random. Table 4a shows the best and worst expected values for objects of various sizes, along with different metrics MSE, MAE, MSLE, and MAPE as given in Table 4b.

From Table 4, it is seen that the ANN model performs exceedingly well on the new data, achieving a MAPE of 1.57%. It achieves a MSE of 0.565 and MAE of 0.383. This time, it achieves a 0% error on couple of occasions when

**Table 4**

Performance of the ANN model on the new data

(a) Predicted values					(b) Performance metrics	
Sl. No.	True values (mm)	Predicted values (mm)		Median (mm)	Metrics	Value
		Worst	Best			
		1	10.0		09.72	10.00
2	15.0	16.06	14.99	MAE	0.383	
3	18.0	17.33	17.98	MAPE	1.567	
4	22.0	23.58	21.96	MSLE	0.001	
5	25.0	24.51	24.98			
6	29.0	24.97	28.86			
7	30.0	28.66	30.00			
8	35.0	36.13	34.96			
9	38.0	37.38	38.01			
10	41.0	40.18	41.14			
11	50.0	47.71	49.97			

**Table 5**

Performance of the ANN model on for metal targets

(a) Predicted values					(b) Performance metrics	
Sl. No.	True values (mm)	Predicted values (mm)		Median (mm)	Metrics	Value
		Worst	Best			
		1	10.0		09.72	10.00
2	15.0	16.06	14.99	MAE	0.234	
3	18.0	17.33	17.98	MAPE	1.179	
4	22.0	23.58	21.96	MSLE	0.0003	
5	25.0	24.51	24.98			
6	30.0	28.66	30.00			
7	38.0	37.38	38.01			
8	50.0	47.71	49.97			

estimating the size of objects having diameters 10 mm and 30 mm.

### 3.1.1. Metallic targets

Of the collected 191 A-Scans, the data is segregated into metallic and non-metallic objects and the model's performance is evaluated again. The predictions made by the ANN model for metallic objects are shown in Table 5a. From Table 5b, it is seen that the model is able to achieve a MAPE of 1.18% on metallic targets.

### 3.1.2. Non-metallic targets

Performance is also evaluated for PVC targets. The true and predicted values by the model for non-metallic targets are given in Table 6a. Its performance is slightly poorer than the metallic targets, achieving a MAPE of 3.03%. The signals scattered from non-metal pipes are smaller. The dielectric contrast between the target (pipe) and medium (soil) also affects the GPR signal. The lower the contrast, the weaker the reflected signal. Moreover, this phenomenon is frequency-dependent (Shaari et al., 2010) due to which the model has different MAPE values for metallic and non-metallic targets.

**Table 6**

Performance of the ANN model on for non-metal targets

Sl. No.	True values (mm)	(a) Predicted values		Median (mm)	(b) Performance metrics	
		Predicted values (mm)			Metrics	Value
		Worst	Best			
1	29.0	24.97	28.86	27.6	MSE	2.138
2	35.0	36.13	34.96	34.89	MAE	0.945
3	41.0	40.18	41.14	41.14	MAPE	3.032
					MSLE	0.003

**Table 7**

Comparison of present work with past studies

Sl. No.	Author	Technique used	Application	Accuracy / Error
1	Rathod et al. (2019)	GPR and Profoscope data	Rebar detection and diameter estimation	2-11% error. 100% error for small diameters
2	Jazayeri et al. (2019)	full-waveform inversion of GPR data	Reinforced concrete mapping and rebar diameter estimation	11% error
3	Lei et al. (2020)	Hyperbolic feature extraction CNN-LSTM Framework	Diameter identification of cylindrical objects in GPR B-Scans	92.5% accuracy
4	Giannakis et al. (2021)	Autoencoder as a pre processing step followed by ANN based regressor	Estimate the diameter of reinforcement bars in concrete	$\pm 6$ mm accuracy
5	Barkataki et al. (2022)	Hyperbolic feature detection followed by deep CNN	Estimation of radius/ diameter of buried objects	7% error
6	<b>Present work</b>	Feature extraction using ANN on spectral data of GPR A-Scans	Estimation of diameter of buried objects	1.89% error

## 4. Summary

A new approach to estimate the size of buried objects from GPR A-Scans was proposed in this study. The proposed ANN model is able to extract important features through several layers of neurons to finally estimate the diameter of buried objects with a high degree of accuracy. The overall performance of the ANN model is given in Tables 3 and 4. It was seen that the dataset containing spectrum information of GPR A-Scans outperformed that containing data in the time domain. The proposed model's MAPE of 1.89% on the validation set, and 1.57% on new unseen data demonstrates its high reliability and good repeatability. It is to be noted that the highest errors were found for non-metallic targets for smaller diameter values.

A comparison with previously reported techniques is shown in Table 7 from which it can be concluded that the use of spectrum information along with the application of ANN in this problem has demonstrated an improved performance. The current work can be extended to estimating sizes of other anomalies of the subsurface like voids in concrete, tunnels etc.

## Acknowledgements

The authors extend their gratitude to several individuals for their invaluable contributions to this study. Firstly, Dr. Manoj Kumar Phukan (Principal Scientist, Geo Sciences and Technology Division, CSIR-NEIST, Jorhat) provided insightful guidance on GPR data interpretation. Ms. Sharmistha Mazumdar's assistance in visualizing the results was also greatly appreciated. Additionally, Mr. Anil Kumar Rajak, Mr. Hafizuddin Ahmed, Mr. Hriday Jyoti Saikia, and Mr. Tilak Chandra Deka provided valuable support during the construction of the experimental setup. The authors

would also like to acknowledge Mr. Ankur Jyoti Kalita for his contributions to the database creation process. Lastly, the authors express their gratitude to Ms. Sweta Kumar Gaur for her suggestions that helped to polish the manuscript linguistically.

## Data Availability

The datasets generated during and/or analysed during the current study are available from the corresponding author on reasonable request.

## CRedit authorship contribution statement

**Nairit Barkataki:** Conceptualization, Methodology, Formal analysis and investigation, Writing - Original Draft. **Banty Tiru:** Supervision, Writing - Review & Editing. **Utpal Sarma:** Conceptualisation, Resources, Supervision, Writing - Review & Editing.

## References

- Ahmed, A.M., Deo, R.C., Ghahramani, A., Raj, N., Feng, Q., Yin, Z., Yang, L., 2021. Lstm integrated with boruta-random forest optimiser for soil moisture estimation under rcp4.5 and rcp8.5 global warming scenarios. *Stochastic Environmental Research and Risk Assessment*, 1–31.
- Akinsunmade, A., Tomecka-Suchoń, S., Pysz, P., 2019. Correlation between agrotechnical properties of selected soil types and corresponding gpr response. *Acta Geophysica* 67, 1913–1919.
- Ayala-Cabrera, D., Izquierdo, J., 2021. Gpr image interpretation advancement for smarter technical management of water leakage in urban water infrastructures. in: *Earth Resources and Environmental Remote Sensing/GIS Applications XII, SPIE*. pp. 220–227.
- Barkataki, N., Borah, P., Sarma, U., Tiru, B., 2021. Design of a 400 mhz cavity backed cpw fed bow-tie antenna for gpr applications, in: *2021 International Conference on Industrial Electronics Research and Applications (ICIERA)*, IEEE. pp. 1–6.
- Barkataki, N., Tiru, B., Sarma, U., 2022. A cnn model for predicting size of buried objects from gpr b-scans. *Journal of Applied Geophysics* 200, 104620.
- Barrile, V., Pucinotti, R., 2005. Application of radar technology to reinforced concrete structures: a case study. *NDT & e International* 38, 596–604.
- Benedetto, A., Pajewski, L., 2015. *Civil engineering applications of ground penetrating radar*. Springer.
- Bi, W., Zhao, Y., An, C., Hu, S., 2018. Clutter elimination and random-noise denoising of gpr signals using an svd method based on the hankel matrix in the local frequency domain. *Sensors* 18, 3422.
- Che, C.K.N.A.H., Melor, K., Joret, A., Razali, M., Ponniran, A., Sulong, M.S., Omar, R., 2021. Frequency based signal processing technique for pulse modulation ground penetrating radar system. *International Journal of Electrical and Computer Engineering* 11, 4104.
- Costamagna, E., Gamba, P., Lossani, S., 1998. A neural network approach to the interpretation of ground penetrating radar data, in: *IGARSS'98. Sensing and Managing the Environment. 1998 IEEE International Geoscience and Remote Sensing Symposium Proceedings*.(Cat. No. 98CH36174), IEEE. pp. 412–414.
- Daniels, D.J., 2004. *Ground penetrating radar*. The Institution of Electrical Engineers.
- Gamba, P., Lossani, S., 2000. Neural detection of pipe signatures in ground penetrating radar images. *IEEE Transactions on Geoscience and Remote Sensing* 38, 790–797.
- Giannakis, I., Giannopoulos, A., Warren, C., 2021. A machine learning scheme for estimating the diameter of reinforcing bars using ground penetrating radar. *IEEE Geoscience and Remote Sensing Letters* 18, 461–465.
- Hasan, M.I., Yazdani, N., 2016a. An experimental and numerical study on embedded rebar diameter in concrete using ground penetrating radar. *Chin. J. Eng* 2016, 1–7.
- Hasan, M.I., Yazdani, N., 2016b. An experimental study for quantitative estimation of rebar corrosion in concrete using ground penetrating radar. *Journal of Engineering* 2016.
- He, W., Hao, T., Ke, H., 2021. Frequency 3d slice image visualization for gpr applications, in: *IOP Conference Series: Earth and Environmental Science*, IOP Publishing. p. 012019.
- Jazayeri, S., Kruse, S., Hasan, I., Yazdani, N., 2019. Reinforced concrete mapping using full-waveform inversion of gpr data. *Construction and Building Materials* 229, 117102.
- Jol, H.M., 2008. *Ground penetrating radar theory and applications*. Elsevier.
- Kalogeropoulos, A., Hugenschmidt, J., van der Kruk, J., Bikowski, J., Brühwiler, E., 2012. Gpr full-waveform inversion of chloride gradients in concrete, in: *2012 14th International Conference on Ground Penetrating Radar (GPR)*, IEEE. pp. 320–323.
- Kalogeropoulos, A., Van der Kruk, J., Hugenschmidt, J., Busch, S., Merz, K., 2011. Chlorides and moisture assessment in concrete by gpr full waveform inversion. *Near Surface Geophysics* 9, 277–286.
- Kang, S., Yu, J.D., Han, W., Lee, J.S., 2022. Nondestructive detection of cavities beneath concrete plates using ground penetrating radar and microphone. *NDT & E International*, 102663.
- Ko, K.H., Jang, G., Park, K., Kim, K., 2012. Gpr-based landmine detection and identification using multiple features. *International journal of antennas and propagation* 2012.
- Kumar, B.S., Sahoo, A.K., Maiti, S., 2021. Removal of clutter and random noise for gpr images, in: *2021 IEEE 18th India Council International Conference (INDICON)*, IEEE. pp. 1–6.
- Lai, W.L., Kind, T., Stoppel, M., Wiggenshauser, H., 2013. Measurement of accelerated steel corrosion in concrete using ground-penetrating radar and a modified half-cell potential method. *Journal of Infrastructure Systems* 19, 205–220.
- Laurens, S., Balayssac, J., Rhazi, J., Klysz, G., Arliguie, G., 2005. Non-destructive evaluation of concrete moisture by gpr: experimental study and direct modeling. *Materials and structures* 38, 827–832.

- Lei, W., Hou, F., Xi, J., Tan, Q., Xu, M., Jiang, X., Liu, G., Gu, Q., 2019. Automatic hyperbola detection and fitting in gpr b-scan image. *Automation in Construction* 106, 102839.
- Lei, W., Luo, J., Hou, F., Xu, L., Wang, R., Jiang, X., 2020. Underground cylindrical objects detection and diameter identification in gpr b-scans via the cnn-lstm framework. *Electronics* 9, 1804.
- Liu, H., Lin, C., Cui, J., Fan, L., Xie, X., Spencer, B.F., 2020. Detection and localization of rebar in concrete by deep learning using ground penetrating radar. *Automation in construction* 118, 103279.
- Liu, S., Li, Z., Zhao, G., 2021. Attenuation characteristics of ground penetrating radar electromagnetic wave in aeration zone. *Earth Science Informatics* 14, 259–266.
- Luo, X., Bai, X., Guo, S., Wang, L., Mi, H., Chen, H., 2020. Machine learning methods applied in detection of buried targets for ground penetrating radar, in: *IET International Radar Conference (IET IRC 2020)*, IET. pp. 450–454.
- Mechbal, Z., Khamlichi, A., 2017. Determination of concrete rebars characteristics by enhanced post-processing of gpr scan raw data. *NDT & E International* 89, 30–39.
- Mertens, L., Persico, R., Matera, L., Lambot, S., 2015. Automated detection of reflection hyperbolas in complex gpr images with no a priori knowledge on the medium. *IEEE Transactions on Geoscience and Remote Sensing* 54, 580–596.
- Miccinesi, L., Beni, A., Monchetti, S., Betti, M., Borri, C., Pieraccini, M., 2021. Ground penetrating radar survey of the floor of the accademia gallery (florence, italy). *Remote Sensing* 13, 1273.
- Pajewski, L., Fontul, S., Solla, M., 2019. Ground-penetrating radar for the evaluation and monitoring of transport infrastructures, in: *Innovation in Near-Surface Geophysics*. Elsevier, pp. 341–398.
- Park, S., Kim, J., Jeon, K., Kim, J., Park, S., 2021. Improvement of gpr-based rebar diameter estimation using yolo-v3. *Remote Sensing* 13, 2011.
- Pasculli, D., Natali, A., Salvatore, W., Morelli, F., Morandi, D., 2018. Investigation of reinforced concrete bridges by using a dual-polarized high-frequency gpr, in: *2018 17th International Conference on Ground Penetrating Radar (GPR)*, IEEE. pp. 1–5.
- Persico, R., 2014. *Introduction to ground penetrating radar: inverse scattering and data processing*. John Wiley & Sons.
- Rasol, M., Pais, J.C., Pérez-Gracia, V., Solla, M., Fernandes, F.M., Fontul, S., Ayala-Cabrera, D., Schmidt, F., Assadollahi, H., 2022. Gpr monitoring for road transport infrastructure: A systematic review and machine learning insights. *Construction and Building Materials* 324, 126686.
- Rathod, H., Debeck, S., Gupta, R., Chow, B., 2019. Applicability of gpr and a rebar detector to obtain rebar information of existing concrete structures. *Case Studies in Construction Materials* 11, e00240.
- Rodés, J.P., Pérez-Gracia, V., Martínez-Reguero, A., 2015. Evaluation of the gpr frequency spectra in asphalt pavement assessment. *Construction and Building Materials* 96, 181–188.
- Salinas Naval, V., Santos-Assunção, S., Pérez-Gracia, V., 2018. Gpr clutter amplitude processing to detect shallow geological targets. *Remote Sensing* 10, 88.
- Santos, V.R.N.d., Al-Nuaimy, W., Porsani, J.L., Hirata, N.S.T., Alzubi, H.S., 2014. Spectral analysis of ground penetrating radar signals in concrete, metallic and plastic targets. *Journal of Applied Geophysics* 100, 32–43.
- Shaari, A., Ahmad, R., Chew, T., 2010. Effects of antenna-target polarization and target-medium dielectric contrast on gpr signal from non-metal pipes using fdtd simulation. *NDT & E International* 43, 403–408.
- Shaw, M., Millard, S., Molyneaux, T., Taylor, M., Bungey, J., 2005. Location of steel reinforcement in concrete using ground penetrating radar and neural networks. *Ndt & E International* 38, 203–212.
- Singh, N.P., Nene, M.J., 2013. Buried object detection and analysis of gpr images: Using neural network and curve fitting, in: *2013 Annual International Conference on Emerging Research Areas and 2013 International Conference on Microelectronics, Communications and Renewable Energy*, IEEE. pp. 1–6.
- Solla, M., Pérez-Gracia, V., Fontul, S., 2021. A review of gpr application on transport infrastructures: Troubleshooting and best practices. *Remote Sensing* 13, 672.
- Sun, H.H., Lee, Y.H., Luo, W., Ow, L.F., Yusof, M.L.M., Yucel, A.C., 2021. Compact dual-polarized vivaldi antenna with high gain and high polarization purity for gpr applications. *Sensors* 21, 503.
- Tomecka-Suchoń, S., 2019. Ground penetrating radar use in flood prevention. *Acta Geophysica* 67, 1955–1965.
- Tong, Z., Gao, J., Yuan, D., 2020. Advances of deep learning applications in ground-penetrating radar: A survey. *Construction and Building Materials* 258, 120371.
- Torrione, P.A., Morton, K.D., Sakaguchi, R., Collins, L.M., 2013. Histograms of oriented gradients for landmine detection in ground-penetrating radar data. *IEEE transactions on geoscience and remote sensing* 52, 1539–1550.
- Wahab, W.A., Jaafar, J., Yassin, I.M., Ibrahim, M.R., 2013. Interpretation of ground penetrating radar (gpr) image for detecting and estimating buried pipes and cables, in: *2013 IEEE International Conference on Control System, Computing and Engineering*, IEEE. pp. 361–364.
- Wang, Y., Cui, G., Xu, J., 2020. Semi-automatic detection of buried rebar in gpr data using a genetic algorithm. *Automation in Construction* 114, 103186.
- Wu, K., Rodríguez, G.A., Zajc, M., Jacquemin, E., Clément, M., De Coster, A., Lambot, S., 2019. A new drone-borne gpr for soil moisture mapping. *Remote Sensing of Environment* 235, 111456.
- Zhang, X., Xue, F., Wang, Z., Wen, J., Guan, C., Wang, F., Han, L., Ying, N., 2021. A novel method of hyperbola recognition in ground penetrating radar (gpr) b-scan image for tree roots detection. *Forests* 12, 1019.
- Zhao, S., Al-Qadi, I., 2017. Pavement drainage pipe condition assessment by gpr image reconstruction using fdtd modeling. *Construction and Building Materials* 154, 1283–1293.
- Zhou, F., Chen, Z., Liu, H., Cui, J., Spencer, B.F., Fang, G., 2018a. Simultaneous estimation of rebar diameter and cover thickness by a gpr-emi dual sensor. *Sensors* 18, 2969.
- Zhou, X., Chen, H., Li, J., 2018b. An automatic gpr b-scan image interpreting model. *IEEE Transactions on Geoscience and Remote Sensing* 56, 3398–3412.