

REPUBLIC OF TURKEY
YILDIZ TECHNICAL UNIVERSITY
GRADUATE SCHOOL OF SCIENCE AND ENGINEERING

**GROUND PENETRATING RADAR ANTENNA DESIGN TO
DETECT BURIED OBJECT AND SIGNAL PROCESSING
WITH DEEP LEARNING NETWORKS BY USING
NUMERICAL ELECTROMAGNETIC METHODS**

Reyhan YURT

DOCTOR OF PHILOSOPHY THESIS
Department of Electronics and Communication Engineering
Communication Program

Supervisor
Assoc. Prof. Dr. Hamid TORPİ

May, 2023

REPUBLIC OF TURKEY
YILDIZ TECHNICAL UNIVERSITY
GRADUATE SCHOOL OF SCIENCE AND ENGINEERING

**GROUND PENETRATING RADAR ANTENNA DESIGN TO DETECT
BURIED OBJECT AND SIGNAL PROCESSING WITH DEEP
LEARNING NETWORKS BY USING NUMERICAL
ELECTROMAGNETIC METHODS**

A thesis submitted by Reyhan YURT in partial fulfillment of the requirements for the degree of **DOCTOR OF PHILOSOPHY** is approved by the committee on 05.05.2023 in Department of Electronics and Communication Engineering, Communication Program.

Assoc. Prof. Dr. Hamid TORPİ
Yildiz Technical University
Supervisor

Approved By the Examining Committee

Assoc. Prof. Dr. Hamid TORPİ, Supervisor
Yildiz Technical University

Prof. Dr. Ahmet KIZILAY, Member
Yildiz Technical University

Asst. Prof. Dr. Alper ÇALIŞKAN, Member
Yildiz Technical University

Prof. Dr. Mustafa Emre AYDEMİR, Member
İstanbul Esenyurt University

Prof. Dr. Sedef KENT PINAR, Member
İstanbul Technical University

I hereby declare that I have obtained the required legal permissions during data collection and exploitation procedures, that I have made the in-text citations and cited the references properly, that I haven't falsified and/or fabricated research data and results of the study and that I have abided by the principles of the scientific research and ethics during my Thesis Study under the title of Ground Penetrating Radar Antenna Design to Detect Buried Object and Signal Processing with Deep Learning Networks by Using Numerical Electromagnetic Methods supervised by my supervisor, Assoc. Prof. Dr. Hamid TORPİ. In the case of a discovery of false statement, I am to acknowledge any legal consequence.

Reyhan YURT

Signature



This study was supported by the Scientific and Technological Research Council of Turkey (TUBITAK) Grant No: 2211-A and the project number of 119N196.

To my family...



ACKNOWLEDGEMENTS

I would first like to thank to my esteemed advisor Assoc. Prof. Dr. Hamid TORPİ for his time, guidance, suggestions, contributions and valuable supports throughout my thesis study. I would like to thank to the members of my thesis committee, Prof. Dr. Ahmet KIZILAY and Prof. Dr. Mustafa Emre AYDEMİR for their useful comments, suggestions and sharing their experiences and knowledge with me during thesis study. I am also grateful to other examining committee members Prof. Dr. Sedef KENT PINAR and Asst. Prof. Dr. Alper ÇALIŞKAN for reading the manuscript, their advices and valuable comments on the thesis.

I would like to thank Scientific and Technological Research Council of Turkey (TUBİTAK) for supporting me with 2211-A doctoral scholarship during my PhD education and for supporting my thesis study with project number of 119N196.

I would like to thank to project manager Prof. Dr. Ahmet KIZILAY once again for providing me the opportunity to get involved in his project, always guiding me with his valuable information and Prof. Dr. Ahmet Serdar TÜRK for sharing his experiences and knowledge, also all project members for their valuable contributions to my thesis study. I would like to offer special thanks to Assoc. Prof. Dr. Peyman MAHOUTİ. I am really thankful to him for his help, encouragement, sharing knowledge on multiple subjects and valuable contributions to my studies.

I would like to thank to my family for their support, unfailing belief and encouragement throughout my studies. I would also like to thank to my elder brother Erhan YURT for his valuable contributions to the preparation of the visuals of my thesis. I offer my special thanks and my endless gratitude to my family for their love, helping me with their best wishes. I am grateful to them for always being there for me, endless love and support.

Reyhan YURT

TABLE OF CONTENTS

LIST OF SYMBOLS	ix
LIST OF ABBREVIATIONS	xi
LIST OF FIGURES	xiii
LIST OF TABLES	xix
ABSTRACT	xxii
ÖZET	xxiv
1 INTRODUCTION	1
1.1 Literature Review	1
1.2 Objective of the Thesis	4
1.3 Hypothesis	4
2 BURIED OBJECT CHARACTERIZATION WITH A-SCAN ANALYSIS AND DATA-DRIVEN SURROGATE MODELS	5
2.1 Introduction to Buried Object Characterization via A-scan Analysis . . .	5
2.2 The Problem Formulation of Data-Driven Surrogate Models	6
2.3 Configuration of GPR Model and Design Scenarios	6
2.4 Generated Data Sets for Buried Object Characterization	9
2.4.1 Arrangement of Noise Free Data Sets	10
2.4.2 Generation of Noisy Data Sets for Supplementary Evaluation .	11
2.5 Data-Driven Surrogate Models	14
2.5.1 State-of-the-Art of Surrogate-Based Buried Object Characterization	15
2.5.2 Proposed Time-Frequency Regression Model Technique for Surrogate Modeling Buried Object	16
2.6 B-scan Evaluation	19
2.7 Validation of Data-Driven Surrogate Modeling Using Measurement Data	24
2.8 Conclusion of Buried Object Characterization with A-scan Analysis and Data-Driven Surrogate Models	28

3	BURIED OBJECT CHARACTERIZATION VIA REGRESSION-ENABLED HYPERBOLIC SIGNATURE EXTRACTION AND DATA-DRIVEN SURROGATE MODELS	30
3.1	Introduction to Buried Object Characterization via Regression-Enabled Hyperbolic Signature Extraction & Problem Formulation	30
3.2	Configuration of GPR Model	31
3.3	Arrangement of Data Sets	37
3.4	Linear Regression-Assisted Surrogate Modeling for Buried Object Characterization	37
3.4.1	Hyperbolic Signature Extraction by Linear Regression Algorithm	38
3.4.2	Deep Learning Based Modified Multilayer Perceptron (M2LP) Framework	41
3.5	Design of Experiments and Results	43
3.5.1	Modeling Results	46
3.5.2	Benchmarking	48
3.6	Data-Driven Surrogate Modeling with Noisy Data for Validation of Realistic Scenarios	54
3.7	Data-Driven Surrogate Models for Buried Object Characterization via Measurement Data	56
3.8	Conclusion of Buried Object Characterization via Data-Driven Surrogates and Regression-Enabled Hyperbolic Signature Extraction .	59
4	COMPERHENSIVE ANALYSIS OF BURIED OBJECT CHARACTERIZATION WITH SEVERAL VARIED DATA SETS AND DATA-DRIVEN SURROGATE MODELS	62
4.1	Introduction to Characterization Analysis of Buried Object via Several Varied Data Sets and Data-Driven Surrogate Models	62
4.2	Configuration of GPR Model and Scenarios of Data Acquisition	63
4.3	Study Cases for Characterization Analysis	66
4.3.1	Case 1- Data Set of Raw Time Signals in 2D	66
4.3.2	Case 2- Data Set of Extracted Feature Based on Raw Time Signals	66
4.3.3	Case 3- Data Set of Consecutive Raw A-scans (Time Signals in 1D)	68
4.4	Surrogate Modeling of Buried Object Characterization via Deep Regression Network (DRN) Framework Customized for Time Frequency Spectrogram (TFS)	70
4.5	Benchmark Models	72
4.6	The Results of Surrogate Models with Study Cases	73
4.7	Characterization Analysis for Validation of Case 3 Under Realistic Scenarios with Noisy Data	78

5 RESULTS AND DISCUSSION	81
REFERENCES	83
PUBLICATIONS FROM THE THESIS	89



LIST OF SYMBOLS

c	Speed of Ligth
D	Depth of Buried Object
d_a	Distance between the Antennas
dB	Decibel
d_r	Distance from the Object Center to the Receiver Antenna
d_t	Distance from the Object Center to the Transmitter Antenna
f_{BW}	Bandwidth
f_c	Center Frequency
ID	Identification Number of A-scan
k	Position Number at the Scanning Aperture
N	Total Number of Samples
N_s	Frequency Step Number
N_t	Time Steps
P	Lateral Position
P_i	i th Model-Predicted Value
R	Radius
t	Time
T	Time Window
t_0	Relaxation Time
t_d	Traveling Time
T_i	i th Target Value
Δf	Frequency Bin Spacing
ϵ_0	Permittivity of Free Space

ϵ_{∞}	Relative Permittivity Value at Infinity Frequency
ϵ_r	Permittivity
ϵ_s	Relative Permittivity Value at Zero Frequency
λ	Wavelength
σ	Conductivity



LIST OF ABBREVIATIONS

AI	Artificial Intelligence
BN	Batch Normalization
C3	Column-Connection Clustering
CNN	Convolutional Neural Network
DL	Deep Learning
DRN	Deep Regression Network
DS	Data Set
EEMD	Ensemble Empirical Mode Decomposition
EM	Electromagnetic
FC	Fully Connected
FDTD	Finite Difference Time Domain
FFT	Fast Fourier Transform
GAs	Genetic Algorithms
K-SVD	K-Singular Value Decomposition
LHS	Latin Hypercube Sampling
M2LP	Modified Multilayer Perceptron
MAE	Mean Absolute Error
MLP	Multilayer Perceptron
NN	Neural Network
PCA	Principal Component Analysis
PEC	Perfect Electric Conductor
PML	Perfectly Matched Layer
ReLU	Rectified Linear Unit

RF	Random Forest
RME	Relative Mean Error
SNR	Signal-to-Noise Ratio
STFT	Short Time Fourier Transform
SVM	Support Vector Machine
SVRM	Support Vector Regression Machine
TFRM	Time-Frequency Regression Model
TFS	Time Frequency Spectrogram



LIST OF FIGURES

Figure 2.1	Diagram explaining the geometry of the GPR model, the problem formulation and the characteristic parameters of the buried object to be identified	7
Figure 2.2	The 3D view of the GPR model used for generating train and test data sets	8
Figure 2.3	Configuration of characterization parameters in 3D plot in terms of depth, lateral position and radius for scenarios in the training and test data sets	10
Figure 2.4	Black box representation of the proposed data driven surrogate model for buried object characterization	11
Figure 2.5	(a) Input signal and a sample of raw A-scan signals for water content of 2.8 %, depth 250mm, lateral position 210mm, and radius 40mm (b) B-scan image construction sample for the same scenario	13
Figure 2.6	Samples of noisy A-scan signals obtained for SNR value of 30dB and 20dB. (a) A-scan samples of SNR 30dB for a scenario of water content of 2.8 %, depth 250mm, lateral position 210mm, and radius 40mm. (b) A-scan samples of SNR 20dB for a scenario of water content of 5.5 %, depth 150mm, lateral position 140mm, and radius 30mm	14
Figure 2.7	The obtained 32x32 time-frequency images for different signals . .	18
Figure 2.8	Architecture of the proposed TFRM model	19
Figure 2.9	The training loss of the proposed TFRM framework versus the iteration number during the all epochs (epoch number equals to 300)	20

Figure 2.10	Inversion results of surrogate modeling with TFRM and 1-D CNN for B-scan evaluation of raw time signals, the image where the dark gray shows target and light gray predicted in the geometry of subsurface, (a) ($D = 137$ mm, $P = 146$ mm, $R = 25$ mm) estimated by TFRM, (b) ($D = 137$ mm, $P = 146$ mm, $R = 25$ mm) estimated by CNN, (c) ($D = 259$ mm, $P = 163$ mm, $R = 18$ mm) estimated by TFRM, (d) ($D = 259$ mm, $P = 163$ mm, $R = 18$ mm) estimated by CNN (e) ($D = 147$ mm, $P = 193$ mm, $R = 42$ mm) estimated by TFRM, (f) ($D = 147$ mm, $P = 193$ mm, $R = 42$ mm) estimated by CNN	22
Figure 2.11	Inversion results of surrogate modeling with TFRM and 1-D CNN for B-scan evaluation of noisy raw time signals, the image where the dark gray shows target and light gray predicted in the geometry of subsurface, (a) ($D = 259$ mm, $P = 163$ mm, $R = 18$ mm) estimated by TFRM [30 dB], (b) ($D = 259$ mm, $P = 163$ mm, $R = 18$ mm) estimated by CNN [30 dB], (c) ($D = 204$ mm, $P = 241$ mm, $R = 27$ mm) estimated by TFRM [20 dB], (d) ($D = 204$ mm, $P = 241$ mm, $R = 27$ mm) estimated by CNN [20 dB]	25
Figure 2.12	The picture of the experimental environment and setup used to generate the measurement data set	26
Figure 2.13	Samples of measured A-scan signals obtained from the sand pool by using the experimental set up. A-scan samples belong to a scenario given by the object depth 115mm, lateral position 550mm, and radius 25mm	27
Figure 2.14	Inversion results of surrogate modeling with TFRM and 1-D CNN for B-scan evaluation of hold-out measurement data, the image where the dark gray shows target and light gray predicted in the geometry of subsurface, (a) ($D = 110$ mm, $P = 770$ mm, $R = 20$ mm) estimated by TFRM, (b) ($D = 110$ mm, $P = 770$ mm, $R = 20$ mm) estimated by CNN, (c) ($D = 165$ mm, $P = 900$ mm, $R = 15$ mm) estimated by TFRM, (d) ($D = 165$ mm, $P = 900$ mm, $R = 15$ mm) estimated by CNN	27
Figure 3.1	Buried object characterization problem: graphical illustration and definition of terms. Variables R , D , and P stand for the object radius, its depth, and lateral positions, respectively, all to be identified in the characterization process	31

Figure 3.2	Configuration of the GPR model for generating training and testing data used by the proposed surrogate-assisted buried object characterization framework. T and R stand for the transmitter and receiver antenna, respectively. As mentioned earlier, the parameters to be estimated are the lateral position P , depth D , and object radius R	33
Figure 3.3	Signals obtained from the GPR model: (a) exemplary samples of raw A-scan signals from the test object set at $D = 177$ mm, $P = 175$ mm, and $R = 19$ mm; (b) B-scan image construction for the same test object	34
Figure 3.4	Reflected Signals from air-ground boundary and the objects in different depths with their enveloped forms: (a) the received direct wave in the air and reflection from the air-ground boundary; (b) sample scenarios of the A-scan to obtain soil attenuation along the penetration path with the object set at $D = 40$ mm, 80 mm, 200 mm, 520 mm, 560 mm, 570 mm, and $P = 238$ mm as well as $R = 10$ mm; (c) amplitude change of the reflected signals versus the depth due to soil attenuation; (d) amplitude ratio between the incident signal and the reflected signals versus the depth	36
Figure 3.5	The parameters associated with the hyperbolic pattern constructed by the required time of the first reflections from the buried cylindrical object	41
Figure 3.6	Raw and pre-processed B-scan images with their extracted hyperbolic patterns for selected samples concerning the following two test scenarios. The first scenario corresponding to $D = 276$ mm, $P = 172$ mm, and $R = 33$ mm: (a) raw B-scan image, (b) B-scan image with removed background reflections, (c) extracted hyperbolic pattern. The second scenario corresponding to $D = 233$ mm, $P = 83$ mm, and $R = 34$ mm: (d) raw B-scan image, (e) B-scan image with removed background reflections, (f) extracted hyperbolic pattern	42
Figure 3.7	Modeling framework: (a) inputs and outputs of the phases in linear regression assisted surrogate modeling, (b) flowchart of the proposed data driven surrogate model for buried object characterization	44
Figure 3.8	Architecture of the proposed M2LP framework	45
Figure 3.9	The training history of the proposed M2LP model in terms of RMSE (Root Mean Square Error) versus iteration number	45

Figure 3.10	Prediction of geophysical parameters of the object, obtained using the proposed surrogate model (built based on the DS4 dataset). The target and surrogate-predicted object marked using the dark- and light-grey shade, respectively: (a) $D = 249$ mm, $P = 254$ mm, $R = 38$ mm, (b) $D = 348$ mm, $P = 227$ mm, $R = 27$ mm, (c) $D = 125$ mm, $P = 69$ mm, $R = 40$ mm, (d) $D = 227$ mm, $P = 270$ mm, $R = 23$ mm	47
Figure 3.11	Prediction of characteristic parameters of the object, obtained using the proposed surrogate model and the benchmark techniques. The target and surrogate-predicted object marked using the dark- and light-grey shade, respectively: Scenario I: $D = 193$ mm, $P = 110$ mm, $R = 34$ mm, (a) proposed model, (b) CNN; Scenario II: $D = 361$ mm, $P = 189$ mm, $R = 21$ mm, (c) proposed model, (d) MLP; Scenario III: $D = 244$ mm, $P = 299$ mm, $R = 26$ mm, (e) proposed model, (f) SVRM	51
Figure 3.12	A sample test scenario corresponding to $D = 276$ mm, $P = 172$ mm, and $R = 33$ mm: (a) extracted principle component features by using PCA; (b) pre-processed (clutter reduced) B-scan image by using PCA	53
Figure 3.13	Raw and pre-processed B-scan data along with their extracted hyperbolic patterns for selected samples from the test scenarios including noise. The SNR for first scenario is 30 dB, whereas the objective parameters are $D = 212$ mm, $P = 318$ mm, and $R = 19$ mm: (a) raw B-scan image, (b) B-scan image with removed background reflections, (c) extracted hyperbolic pattern. The SNR for the second scenario is 20 dB, $D = 259$ mm, $P = 231$ mm, and $R = 12$ mm: (d) raw B-scan image, (e) B-scan image with removed background reflections, (f) extracted hyperbolic pattern	55
Figure 3.14	Illustration of the GPR configuration utilized for generating experimental training and testing data used by the proposed surrogate-assisted buried object characterization framework	58
Figure 3.15	Raw and pre-processed B-scan data with their extracted hyperbolic patterns for selected sample concerning the following test scenario which corresponds to $D = 750$ mm, $P = 90$ mm, and $R = 15$ mm: (a) raw B-scan image, (b) B-scan image with removed background reflections, (c) extracted hyperbolic pattern	59

Figure 3.16	Prediction of geophysical parameters of the object, obtained using the proposed surrogate model and the benchmark model, CNN (using measured data). The target and surrogate-predicted object marked using the dark- and light-grey shade, respectively: Scenario I: $D = 90$ mm, $P = 970$ mm, $R = 25$ mm, (a) proposed model, (b) CNN; Scenario II: $D = 150$ mm, $P = 850$ mm, $R = 30$ mm, (c) proposed model, (d) CNN	60
Figure 4.1	The two-dimensional model diagram explaining the geophysical parameters to be solved	64
Figure 4.2	Configuration of characterization parameters for scenarios in the training and testing data sets in 3D plot demonstration with regards to depth, lateral position and radius	66
Figure 4.3	Sample of raw A-scan signals and constructed B-scan images from these signals for the following two test scenarios. Raw A-scans for (a) the first scenario corresponds to water content of 2.8 %, depth 306 mm, lateral position 214 mm, and radius 32 mm and the second scenario corresponds to water content of 5.5 %, depth 218 mm, lateral position 244 mm and radius 27 mm (b) B-scan images for the first and the second scenarios	67
Figure 4.4	Sample principle components of A-scans (A-scan ID: 4, 8, 12 and 16) belonging the scenario corresponds to water content of 5.5%, depth 218 mm, lateral position 244 mm and radius 27 mm; (a) A-scan ID: 4, (b) A-scan ID: 8, (c) A-scan ID: 12 and (d) A-scan ID: 16	69
Figure 4.5	Samples of B-scans in a form of consecutive A-scans from test data set for (a) a scenario of water content of 0.2%, depth 311 mm, lateral position 141 mm, and radius 49 mm, (b) a scenario of water content of 5.5%, depth 218 mm, lateral position 244 mm, and radius 27 mm	69
Figure 4.6	Samples of TFS image of consecutive A-scans from test data set for (a) a scenario of water content of 0.2%, depth 311 mm, lateral position 141 mm, and radius 49 mm, (b) a scenario of water content of 5.5%, depth 218 mm, lateral position 244 mm, and radius 27 mm	71
Figure 4.7	The architecture of the proposed deep regression network framework customized for time frequency spectrogram	71
Figure 4.8	The training loss of the proposed DRN model customized for TFS image versus iteration number	73

Figure 4.9	Prediction of characteristic parameters of the object, obtained using the proposed surrogate model DRN. The target and surrogate-predicted object marked using the dark- and light-grey shade in the geometrical configuration, respectively: (a) $D = 379$ mm, $P = 198$ mm, $R = 16$ mm, (b) $D = 320$ mm, $P = 165$ mm, $R = 34$ mm, (c) $D = 268$ mm, $P = 276$ mm, $R = 50$ mm, (d) $D = 195$ mm, $P = 182$ mm, $R = 24$ mm	76
Figure 4.10	Prediction of characteristic parameters of the object, obtained using the proposed surrogate model DRN and the benchmark models. The target and surrogate-predicted object marked using the dark- and light-grey shade in the geometrical configuration, respectively: Scenario I: $D = 193$ mm, $P = 110$ mm, $R = 34$ mm, (a) proposed model, (b) CNN; Scenario II: $D = 361$ mm, $P = 189$ mm, $R = 21$ mm, (c) proposed model, (d) MLP; Scenario III: $D = 244$ mm, $P = 299$ mm, $R = 26$ mm, (e) proposed model, (f) SVRM; Scenario IV: $D = 193$ mm, $P = 110$ mm, $R = 34$ mm, (g) proposed model, (h) GP- regression	77
Figure 4.11	Samples of noisy B-scans in a form of consecutive A-scans from test data set for the first scenario SNR value of 30 dB scenario of water content of 0.2%, depth 311 mm, lateral position 141 mm, and radius 49 mm; for the second scenario SNR value of 20 dB scenario of water content of 6.2%, depth 211 mm, lateral position 225 mm, and radius 15 mm (a) consecutive A-scans and (b) B-scan images for the first and the second scenarios	80

LIST OF TABLES

Table 2.1	Extended Debye parameters for subsurface dielectric features according to water content [30]	8
Table 2.2	The arrangement of the training data scenarios for the considered buried object characterization problem	9
Table 2.3	Accuracy and the training time of the proposed and the benchmark models, averaged with standard deviation over ten independent runs	19
Table 2.4	Breakdown of the best prediction performance of characteristic parameters	20
Table 2.5	TFRM- and CNN-based prediction of the characteristic parameters in comparison with their true values for the selected test scenarios . .	23
Table 2.6	Prediction performance of considered characteristic parameters for noisy data sets	24
Table 2.7	TFRM- and CNN-based prediction of the characteristic parameters in comparison with their true values for the selected test scenarios for noisy data sets	24
Table 2.8	Accuracy and the training time of the proposed and the benchmark model, CNN. Shown are the averaged error metrics with standard deviation obtained for ten independent runs using the measurement data	28
Table 2.9	Breakdown of the prediction performance of the considered characteristic parameters for hold-out data	28
Table 2.10	TFRM- and CNN-based prediction of the characteristic parameters in comparison with their true values for the selected hold-out scenarios for measurement data set	28
Table 3.1	The arrangement of the training/testing data for the considered GPR problem	45
Table 3.2	Prediction performance of considered characteristic parameters for the proposed M2LP framework	46
Table 3.3	M2LP-based estimation of the characteristic parameters in comparison with their true values for selected test scenarios	48

Table 3.4	Characteristic parameter prediction of the proposed M2LP framework assuming short distance between the object and the ground surface	49
Table 3.5	Performance metrics and the training time of the proposed and the benchmark models, averaged with standard deviation over ten independent runs. The models trained using the dataset DS4 (500 samples)	50
Table 3.6	Breakdown of the best prediction performance of the proposed and the benchmark models, over ten independent runs. The models trained using the dataset DS4 (500 samples)	50
Table 3.7	Characteristic parameter prediction for the proposed and benchmark methods for selected test scenarios	52
Table 3.8	Average performance metrics of the proposed and the benchmark cases and standard deviation, computed over ten independent runs	54
Table 3.9	Prediction accuracy of characteristic parameters of the buried object for all models. Considered scenarios include noisy data with SNR of 30 dB and 20 dB	56
Table 3.10	Predicted characteristic parameters of the buried object versus true values for selected test scenarios	57
Table 3.11	Prediction performance of considered characteristic parameters for M2LP and CNN models by using real data set obtained from measurements	58
Table 4.1	Subsurface dielectric features of Extended Debye model according to water content [30]	65
Table 4.2	The arrangement of the training data for the considered GPR problem	65
Table 4.3	Prediction performances and computational costs of the all cases and surrogate models, averaged with standard deviation over ten independent runs	74
Table 4.4	The best prediction performance of considered characteristic parameters for case 3 and the proposed model and benchmark models	75
Table 4.5	DRN framework customized for TFS image estimation of the characteristic parameters in comparison with their true values for selected test scenarios	78
Table 4.6	Characteristic parameter prediction for the proposed DRN and benchmark methods in comparison with their true values for selected test scenarios	78
Table 4.7	Prediction performances of characteristic parameters of the buried object for DRN and CNN models. Considered scenarios include noisy data with SNR of 30 dB and 20 dB	79

Table 4.8 Characteristic parameter prediction for the DRN and CNN models in comparison with their true values for selected noisy test scenarios . 79



Ground Penetrating Radar Antenna Design to Detect Buried Object and Signal Processing with Deep Learning Networks by Using Numerical Electromagnetic Methods

Reyhan YURT

Department of Electronics and Communication Engineering
Doctor of Philosophy Thesis

Supervisor: Assoc. Prof. Dr. Hamid TORPI

In the thesis, the buried object is characterized with artificial-intelligence-based networks using 3-D full-wave and FDTD-based electromagnetic simulations of defined ground penetrating radar (GPR) models to generate B-scan data. The task is to estimate characteristic parameters of a cylindrical shape, perfectly electric conductor (PEC) object of various radii, buried at different positions independently from each other and simultaneously. The proposed methodologies capitalize on fast and accurate data-driven surrogate models developed for object characterization in terms of its depth, lateral position, and radius. For this purpose, three different data sets and methodologies are taken into consideration. Firstly, a novel deep learning-based framework, time-frequency regression model (TFRM) which focuses on the object characterization in terms of its geophysical parameters with A-scan analysis without clutter reduction by using a few linear sampled train scenarios is proposed. Another methodology is achieved by operating at the level of hyperbolic signatures extracted from the B-scan data through linear regression algorithm, which effectively reduces the dimensionality and the size of data. This methodology relies on reduced of 2D B-scan image to 1D data including variation of reflected electric fields' amplitudes with respect to the scanning aperture. The hyperbolic signatures encode information about the characteristic parameters of the buried object, including depth, lateral position, and radius, all of which can be extracted using proposed methodology. Parametric estimation of the object radius and the location parameters simultaneously

is a challenging problem. The application of processing steps on B-scan profiles incurs high computational costs, which is a limitation of the current methodologies. The metamodel itself is rendered using a novel deep-learning-based modified multilayer perceptron (M2LP) framework. The final methodology is proposed a comprehensive characterization analysis with study cases including diverse data sets for object buried in four different soil media. The characteristic parameters independently from each other and simultaneously are estimated via computationally efficient a novel deep regression network (DRN) by using time frequency spectrogram (TFS) of unprocessed, raw consecutive A-scans. The presented DRN framework customized for TFS without any application of clutter reduction method has prediction performance in a qualitative sense and superior to benchmark models and comparative cases. The presented approaches are favourably benchmarked against the latest developed regression techniques, including multilayer perceptron (MLP), convolutional neural network (CNN), Gaussian process (GP) regression and support vector regression machine (SVRM). Furthermore, the proposed surrogate models are validated via measurement data, which is indicative of suitability of the approaches to handle physical measurements as data sources.

Keywords: Buried object characterization, ground penetrating radar (GPR), surrogate modeling, artificial intelligence, deep learning, A-scan data analysis, linear regression, time frequency spectrogram.

Gömülü Hedef Tespit Etmek İçin Yere Nüfuz Eden Radar Anten Tasarımı Ve Sayısal Elektromanyetik Yöntemler Kullanarak Derin Öğrenme Ağları İle Sinyal İşleme

Reyhan YURT

Elektronik ve Haberleşme Mühendisliği Anabilim Dalı

Doktora Tezi

Danışman: Doç. Dr. Hamid TORPİ

Tezde, gömülü nesne, B-tarama verileri üretmek için tanımlanmış yere nüfuz eden radar modellerinin 3 boyutlu tam dalga ve zamanda sonlu farklar yöntemi tabanlı elektromanyetik simülasyonlarını yapay zeka tabanlı ağlarla karakterize edilir. İş, farklı konumlara gömülmüş, çeşitli yarıçaplardaki mükemmel elektrik iletken, silindirik şeklindeki bir nesnenin karakteristik parametrelerini birbirinden bağımsız ve eş zamanlı olarak tahmin etmektir. Önerilen metodolojiler, derinliği, yatay konumu ve yarıçapı cinsinden nesnenin karakterizasyonu için geliştirilen hızlı ve doğru veri odaklı vekil modellerden yararlanmaktadır. Bu amaçla üç farklı veri seti ve metodolojisi dikkate alınmıştır. İlk olarak, birkaç doğrusal örneklenmiş eğitim senaryosu kullanılarak istenmeyen yankıları azaltmadan A-tarama analizi ile geometrik ve fiziksel parametreleri açısından nesne karakterizasyonuna odaklanan yeni bir derin öğrenmeye dayalı yapı, zaman-frekans regresyon modeli önerilmiştir. Diğer bir metodoloji, doğrusal regresyon algoritması aracılığıyla B-tarama verilerinden çıkarılan hiperbolik işaretler düzeyinde çalışılarak elde edilir, bu da verilerin boyutunu ve büyüklüğünü etkili bir şekilde azaltır. Bu metodoloji, yansıyan elektrik alanlarının genliklerinin tarama açıklığına göre değişimini içeren 2B B tarama görüntüsünden indirgenen 1B verilere dayanır. Hiperbolik işaretler, tümü önerilen metodoloji kullanılarak çıkarılabilen derinlik, yatay konum ve yarıçap dahil olmak üzere gömülü nesnenin karakteristik parametreleri hakkındaki bilgileri kodlar. Nesne yarıçapının ve konum parametrelerinin aynı anda parametrik tahmini zor bir sorundur. B-tarama profillerinde işleme adımlarının uygulanması, mevcut

metodolojilerin bir sınırlaması olan yüksek hesaplama maliyetlerine neden olur. Metamodelin kendisi, yeni bir derin öğrenmeye dayalı değiştirilmiş çok katmanlı algılayıcı yapısı kullanılarak oluşturulur. Son metodoloji, dört farklı toprak ortamına gömülü nesne için çeşitli veri kümelerini içeren vaka incelemeleriyle kapsamlı bir karakterizasyon analizi önerilmiştir. Herhangi bir istenmeyen yankıları azaltma yöntemi uygulanmayan ZFS için özelleştirilmiş sunulan DRA yapısı, niteliksel anlamda tahmin performansına sahiptir ve kıyaslama modellerinden ve karşılaştırmalı vakalardan daha üstündür. Sunulan yaklaşımlar, çok katmanlı algılayıcı (MLP), konvolüsyonel sinir ağı (CNN), Gauss süreci (GP) regresyonu ve destek vektör regresyon makinesi (SVRM) dahil olmak üzere son teknoloji regresyon teknikleriyle uygun bir şekilde karşılaştırılmıştır. Ayrıca, önerilen vekil modeller, fiziksel ölçümleri veri kaynağı olarak ele alma yaklaşımlarının uygunluğunun göstergesi olarak ölçüm verileri aracılığıyla doğrulanmaktadır.

Anahtar Kelimeler: Gömülü nesne karakterizasyonu, yere nüfuz eden radar, vekil modelleme, yapay zeka, derin öğrenme, A-tarama veri analizi, doğrusal regresyon, zaman frekans spektrogramı.

1

INTRODUCTION

1.1 Literature Review

Ground penetrating radar (GPR) have been used as near surface remote sensing method and its working principle is based on electromagnetic wave theory. Over the years, as a nondestructive tool, GPR have a several different applications not only detect buried mines, pipes and wires but also to collect data for geophysical, archeological investigations and planetary explorations [1–3]. For investigation and monitoring underground structures, its main operation is sending electromagnetic time or frequency signal and receiving it back by helping of antennas mostly used as transmitter and receiver. The antenna system including transmitter and receiver is moving along a path or synthetic aperture on upper surface of the ground and scanning underground in a typical GPR configuration. Along the synthetic aperture (scanning axis), the reflected received signal at one point, this signal is named A-scan which is 1D time-variable amplitude signal [1–4]. The concatenated form of A-scans obtained at certain intervals along the scanning path is defined as B-scan (2D data, image) [5–12].

There are several methods utilized for reduction the reflections of the ground surface, and the reflections coming from the scanned underground medium [5]. These are generalized as data (or image) pre-processing. Some of the studies in the literature have reported pre-processing methods based on background subtraction [5, 7, 12] for buried object detection and understanding object-related properties, characterization with regard to localization, object size estimation, material type and etc. via a hyperbolic pattern in the B-scan. Since, a buried object with cylindrical shape such as a rebar, a pipeline, and a wire (energy, optical, or signal cable) constructs to a hyperbolic regression in the B-scan data [6]. Therefore, hyperbolic signatures or patterns in B-scan data are investigated in reflected signals (waves) from the buried object [6–14] to obtain knowledge about the geophysical parameters in terms of depth, lateral position and radius of buried object. This technique is commonly applied in recognition, identification and also characterization of buried object by the help of

numerical, analytical and artificial intelligence methods [15–20].

The extraction of hyperbolic pattern from B-scan and its analysis require some pre-processing methods. The operation of background subtraction and elimination of reflections due to boundary between the air and ground surfaces are the most commonly preferred and effectively used techniques for pre-processing B-scan data [6–20]. These techniques make possible to buried object recognition and identification with related properties of object like vertical and horizontal position, type, shape and object size [6–20]. After pre-processing operations, a deep learning framework which is named as Single Shot Multibox Detector (SSD) [6] is utilized for object detection and localization. In addition to investigation of hyperbolic signature by using pre-processing, column-connection clustering (C3) algorithm [7] are proposed to recognize the region of interest, after that outputs of C3 algorithm are classified with neural network (NN) for identification of hyperbolic signature. Moreover, an algorithm for orthogonal-distance fitting [7] is applied to the identified hyperbola. Another study is analyzed hyperbolic curve identification in the pre-processed B-scan image for object detection and localization with the help of genetic algorithms (GAs) [8], also material type is classified by using support vector machine (SVM) [8].

In a varied machine learning framework including wavelet scattering and SVM in cascaded structure [9], pipeline identification with classification approach, localization and estimation of diameter from reconstructed GPR profiles are proposed. Two pre-processing procedures as de- “wow” and amplitude gain are followed to clarify characteristic of buried target, before reconstruction of GPR profiles by using 1D vertical and horizontal signals in separate data sets which is trained in wavelet scattering networks joined to SVM classifier [9]. In addition, a substituted hyperbolic summation method using hyperbola is developed to obtain focused GPR B-scan image [10]. In another study, a removal of subsurface reflection from the extracted hyperbola in the B-scan images with regard to the amplitude and time vector are used to arrange inputs in a cascaded NN architecture for characterization of the buried object [11] by the help of the Hilbert transform (HT) to obtain enveloped signals. Furthermore, normalized B-scan images are used as inputs for convolutional neural networks after removing air-ground reflections [12].

For buried object recognition, characterization by using B-scan data, deep learning algorithms (DL) are successfully used to develop new network models, particularly convolutional neural network (CNN) frameworks [12–14, 16, 17, 19, 21]. In a framework including CNN and LSTM (long-short term memory) in a cascaded structure, 3D GPR data generated with along and cross axes are analyzed for recognition of buried explosive object and discrimination target or nontarget alarms

[13]. Another study for the detection of a cylindrical object [14], CNN are used together with LSTM network. Also, nine different diameters of objects are classified in the extracted hyperbola regions within B-scan data [14] generated using the gprMax toolbox [22, 23]. The customized deep learning network as CNN framework is applied with SVM classifier instead of softmax layer [21], and this structure is proposed for classification of soil type, material type and object shape. In addition, another methodology is permittivity mapping of the subsurface structures for lining detection [24, 25] using specialized CNN, and deep neural network frameworks. Using these tools, dielectric features in tunnel lining are obtained from the inversion of GPR data, and presented with dielectric images [24, 25]. A counterpart approach is applied to acquire permittivity inversion of geo-structures of buried targets by using deep neural network architectures [26].

In literature, detection is not the only problem considered in the context of buried object characterization. Some studies focus on investigating characteristic properties such as material type and material shape classification, localization, medium dielectric features, and object size classification as well as estimation [6–9, 12, 20, 21, 27, 28]. To solve these tasks, artificial-intelligence (AI)-based surrogate modelling approaches are proposed, particularly including cascaded networks [11, 20, 28]. In an example case study, B-scan images generated by means of gprMax electromagnetic simulator toolbox [22, 23] are pre-processed, and windowing operation is applied on pre-processed B-scan images to extract the most related amplitudes with the object [20]. The extracted amplitudes, the results of material type classification, hyperbola curvature and the depth of the object are used to estimate size of the object via Gaussian process (GP) regression [20]. In another study [28], a framework incorporating a Random Forest (RF) routine and NN regression model are employed to predict the object radius, depth, and the water content of the subsurface, based on the compressed reflected signal. The procedure is an exemplary study on the cascaded networks, with the regression parameters such as depth and water content being independent of each other, and the radius being dependent on the other predicted parameters [28]. In [28, 29], A-scan, time-varying amplitude signals are handled as inputs for buried object characterization for generating 2,000 training samples to obtain satisfactory results. Utilization of the A-scan analysis in the mentioned studies, is aimed at realizing practical processing, and a reduction of the needed computational resources for generating the training data sets.

1.2 Objective of the Thesis

The purpose of this study is to create user-defined data sets for GPR applications and develop novel data-driven surrogate modeling approaches for characterization buried objects in terms of characteristic parameters as depth, lateral position and radius. Although, several studies have been in literature to propose the best approaches, detection target with GPR image is inexplicit yet. Therefore, we aim to provide some contributions to literature for buried object characterization problems and propose new modeling approaches by using artificial neural networks in parametric regression approach. In addition, the main aim of the proposed works is to prioritize computationally efficient surrogate modeling studies and propose novel deep learning-based frameworks that focus on novel methodologies for the object characterization in terms of its geometrical and physical parameters.

1.3 Hypothesis

The mentioned studies consist of detection, identification of buried object in terms of material type, object shape and characterization of object with regards to mostly depth and radius, also by using cascaded network frameworks. It is emphasized that this study provides contributions of ongoing research of different architecture of artificial intelligent networks for characterization of buried object with regards to estimation of characteristic parameters independently from each other, and simultaneously. The user-defined data sets have been created by using 3D full-wave simulator and gprMax FDTD-based electromagnetic simulator toolbox according to defined scenarios. The next chapter of thesis, it is aimed a fast and accurate deep learning-based surrogate modeling by using A-scan analysis. The characteristic parameters in terms of depth, lateral position and radius of buried object can be estimated independently from each other and simultaneously, also independent from dielectric features of subsurface medium. In the following chapter, the methodology of regression-enabled hyperbolic signature extraction is proposed to estimate characteristic parameters (depth, lateral position and radius) via a novel deep learning-based framework. The last two chapters consist of a comprehensive analysis with several study cases and surrogate models for buried object characterization and deep regression network customized for time frequency spectrogram of consecutive raw A-scans is proposed, also results and discussions along with some recommendations for future work.

BURIED OBJECT CHARACTERIZATION WITH A-SCAN ANALYSIS AND DATA-DRIVEN SURROGATE MODELS

2.1 Introduction to Buried Object Characterization via A-scan Analysis

The mentioned techniques in literature review require data obtained using a series of signal and image processing methods for identification and characterization of a buried object. In this section, as one of its contributions, the characterization of buried object is achieved by using data-driven surrogate model constructed with sparse dataset, without removing ground reflections, background subtraction operations, B-scan image processing, and hyperbola investigations (unprocessed). In particular, removing ground reflections or background subtraction operations might be challenging for buried object scenarios that belong to more than one soil type or host medium dielectric features, which changes with the percentage of the water content. Herein, three different soil media with respect to their percentage of the water content [28, 29] are taken into account. The main issue of the proposed work is to develop computationally efficient surrogate modeling via a novel deep learning-based framework that focuses on the object characterization in terms of its geophysical parameters with A-scan analysis (1D time-varying amplitude signal) [28, 29] without any background removal, extraction of hyperbola signature by using a few linear sampled train scenarios. In addition to proposing a novel framework that employs the aforementioned a unique methodological approach, the performance comparison with the state of the art techniques reported in the literature is provided for parametric estimation of buried object characterization.

The target characterization is applied independently based on the subsurface dielectric properties. In addition, a novel deep-learning-based model (time-frequency regression model, TFRM) is proposed to simultaneously predict the depth, the location (the lateral position), and the radius of the object, at low computational cost. A-scan data (1D time signal) obtained from 315 scenarios is analyzed to localize and estimate

the object size by using artificial intelligence (AI) techniques. For each scenario, 16 A-scans are taken and data set consist of reflected normalized power amplitude versus propagation time and A-scan ID (A-scan number according to its scanning point at the synthetic aperture).

For the supplementary evaluation of performance of the proposed model by using noise addition to raw A-scans, a new data sets were generated by adding a random noise to the A-scan signals. Thus, the environmental and internal noise of the GPR system and its effect was analyzed under realistic scenarios. For this purpose, noise incorporation is investigated in terms of the performance of the proposed surrogate modeling methodology by using these noisy data sets. Also, the proposed surrogate modeling approach is validated with measurement data.

2.2 The Problem Formulation of Data-Driven Surrogate Models

The primary problem to be solved in this chapter is estimation of characteristic parameters of buried cylindrical PEC object [4, 6, 9, 28, 29] in different soil media by using A-scan analysis as well as surrogate modelling techniques, specifically, a dedicated TFRM framework, described in detail in Section 2.5. The problem formulation has been illustrated with a diagram and geometry of the GPR model for estimation of identified characteristic parameters in Fig. 2.1. The estimation of the object parameters is realized using computationally efficient surrogate model established using unprocessed, raw data obtained from the GPR model to be discussed in next section. The details concerning the configuration of ground penetrating radar model, as well as the data sets being processed have been elaborated on in Sections 2.3 and 2.4, respectively. Section 2.5 provides the details of surrogate modelling method developed to carry out the estimation process, along with the numerical results and geometrical presentation of estimated parameters. Also, Section 2.6 explains the benchmark network models with their numerical and geometrical representation results.

2.3 Configuration of GPR Model and Design Scenarios

In this work, we consider a ground penetrating radar (GPR) problem of characterizing a buried object, based on 3D full-wave electromagnetic (EM) analysis. Here, a C-band pyramidal horn antenna is used in a monostatic configuration [10]. Figure 2.2 shows the configuration of the data acquisition model for the object characterization. The buried scatterer is defined as perfectly electric conductor (PEC) object such as a wire, a pipe, or a rebar. The travelling time of the wave transmitted by the antenna can be

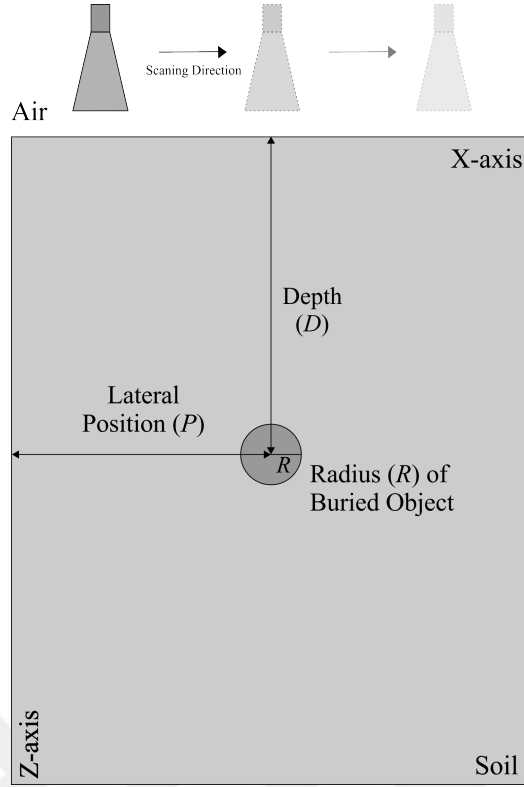


Figure 2.1 Diagram explaining the geometry of the GPR model, the problem formulation and the characteristic parameters of the buried object to be identified

computed using the object depth, the relative permittivity of the subsurface, and the speed of light in the free space. The travelling time t_d of the wave transmitted by the antenna can be computed as

$$t_d = \frac{\sqrt{\epsilon_r} 2d}{c} \quad (2.1)$$

where d is the object depth, ϵ_r is the relative permittivity of the subsurface, and c is the speed of light in the free space. Note that the wave propagation time is monotonically dependent on the depth but also the subsurface permittivity.

The dimensions of scanning subsurface are set to $400 \times 300 \times 500$ [mm]. The surface itself is assumed to be a dispersive soil parameterized by the percentage of water content and the dielectric features, as described by the extended Debye model [30]

$$\epsilon = \epsilon_\infty + \frac{\epsilon_s - \epsilon_\infty}{1 + j\omega t_0} + \frac{\sigma}{j\omega \epsilon_0} \quad (2.2)$$

where ϵ_s and ϵ_∞ are the relative permittivity values at zero and infinite frequency, respectively, t_0 is the relaxation time, σ is the conductivity, w is the angular frequency,

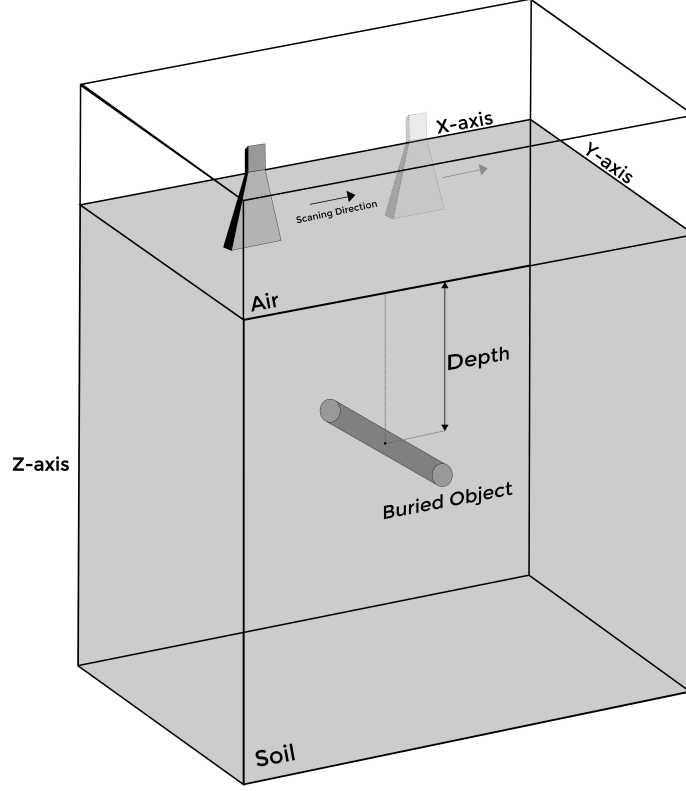


Figure 2.2 The 3D view of the GPR model used for generating train and test data sets

and ϵ_0 is the free space permittivity. These parameters have been demonstrated in Table 2.1 [30] and used to compute the subsurface permittivity. Using this model, three different types of subsurface models are chosen according to their water content of 0.2 %, 2.8 %, and 5.5 % [30]. Also the buried object estimation is carried out without any information of the soil type or dielectric permittivity of the background media. These subsurface models distinguished with respect to the water content, have been chosen similarly as in [28, 29]. GPR scenarios have been configured by placing the aforementioned buried object of various radii at different depths of the dispersive soil medium, and the object has placed at different lateral positions (horizontal positions) based on the origin of scanning path (aperture). Also, 2D geometrical diagram of GPR scenarios for solving the proposed problem has been shown in Fig. 2.1.

Table 2.1 Extended Debye parameters for subsurface dielectric features according to water content [30]

WC %	ϵ_s	ϵ_∞	t_0 [ns]	$\sigma[\Omega^{-1}m^{-1}]$
0.2	4.814	4.507	0.82	6.06×10^{-4}
2.8	6.75	5.503	2.28	2.03×10^{-3}
5.5	8.63	6.023	1	5.15×10^{-3}

2.4 Generated Data Sets for Buried Object Characterization

One of the main contributions of this work is to propose a method for buried cylindrical object [4, 6, 9, 28, 29] characterization, which—to a certain extent—trades off the accuracy for exceptional computationally efficiency. Achieving this purpose requires a construction of the surrogate model using a possibly small and sparse training dataset. The parameter space has four dimensions, which include the buried object position for 2D as depth and lateral position, radius, and the soil water content. Table 2.2 gathers information about the design of experiments, which is a linear sampling method, used for generating 315 sample points within the given variable ranges. For each scenario, 16 A-scans are taken. It should be mentioned that by increasing the number of A-scans for each scenario, it is possible to increase the accuracy of the model for estimating the lateral position and the object size.

However, this would also significantly increase the dataset generation time and computational cost. In this work, the 3D full-wave simulation tool CST Microwave Studio has been used for obtaining the data. For the discretized structure of the GPR model, the average simulation time of each scenario is 14 hours (hardware configuration: Intel(R) Core(TM) i7-CPU @ 2.60GHz Turbo Boost 4.5 GHz, 16 GB RAM). Thus, for the sake of computational efficiency, the number of A-scans has been kept at sixteen. Another comment is that although it is possible to include the water content of the ground as another input variable, such approach might not be feasible in practice due to the inhomogeneities in the examined area. To ensure a more realistic approach, water content of the ground is variable in a discrete sense, i.e., by using three different values. However, this information is not provided to the model neither in training nor the test process. Although this approach would make the modelling process more challenging and be detrimental to the model predictive power, it needs to be followed to be more in line with realistic GPR applications.

Table 2.2 The arrangement of the training data scenarios for the considered buried object characterization problem

Parameter	Sample range	Step size
Depth of Object [mm]	100-400	50
Lateral Position [mm]	140-280	70
Radius of Object [mm]	10-50	10
Water Content[%]	0.2, 2.8 and 5.5	
Number of A-scans in each scenario	16	
Total Number of Scenarios	315	

2.4.1 Arrangement of Noise Free Data Sets

The considered problem of estimating geophysical parameters of a buried object is 2D. The training and testing scenarios include B-scan images (2D data), which contain 16 pairs of A-scan (concatenated forms of A-scans). Each A-scan is a time-varying normalized power amplitude signal obtained at the one point along the synthetic aperture. In other words, the scanning path length is 600. In addition, A-scan combination was presented as the A-scan ID according to points at the scanning path (400 mm). The data consists of 600 time-varying amplitudes and the A-scan ID (601×1). As for the test data set, Latin Hypercube Sampling (LHS) [31] method is used for creating randomly selected sample points to prevent over-fitting of the model. Total of 63 scenarios is generated to be used for performance evaluation of surrogate models. Here, the A-scan ID is an integer between 1 and 16. The training and testing datasets consist of 315 linearly samples scenarios, and 63 randomly selected scenarios, respectively, with the data acquired using a full wave EM simulator. Each data set contains 16 A scan signals (the total of 5040 and 1008 A-scans for training and test data sets respectively). Figure 2.3 presents the configuration of the characterization parameters for the training and test data sets. The considered problem of estimating geophysical parameters of a buried object is 2D.

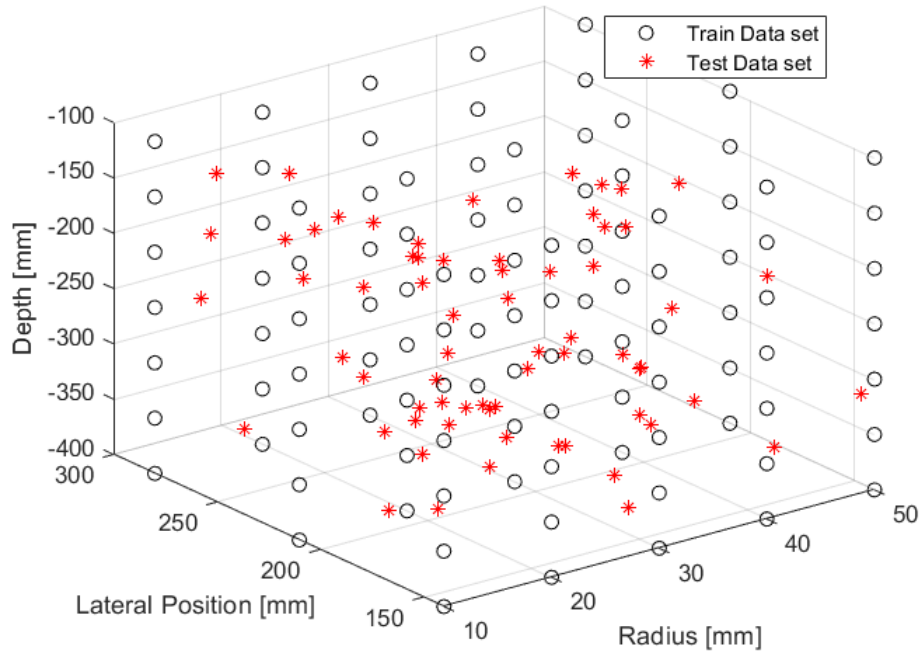


Figure 2.3 Configuration of characterization parameters in 3D plot in terms of depth, lateral position and radius for scenarios in the training and test data sets

In the studied GPR problem, the received reflected signal or statistical features of the signal (power amplitude and A-scan ID) are used as inputs of AI based surrogate

models. The latter are employed to predict the characteristics of the buried objects. A-scans have been used as inputs to enable practical processing, as well as reduction of the necessary computational resources for generating the training/testing datasets [28, 29]. In Fig. 2.4 a black-box representation of the proposed data-driven surrogate modeling approach is presented. 601 features that are defined on the left-hand-side of the black box are the input variables of the problem to be used for the characterization of the targeted outputs (depth, lateral position, and the radius of the object), presented on the right side of the model.

Raw time-variable signal dataset consists of the power amplitude of the received reflected signals versus the time of travelling wave at the penetration axis which is taken as 12 ns using the travelling time, the depth, and the subsurface permittivity relation. Figure 2.5 shows a modulated Gaussian signal with a center frequency of 6 GHz, used as the excitation signal, along with the example of A-scans obtained for a specific scenario. The same figure shows a B-scan image constructed by combining the A-scans obtained as a result of reflections of the transmitting excitation signal along a selected scanning path.

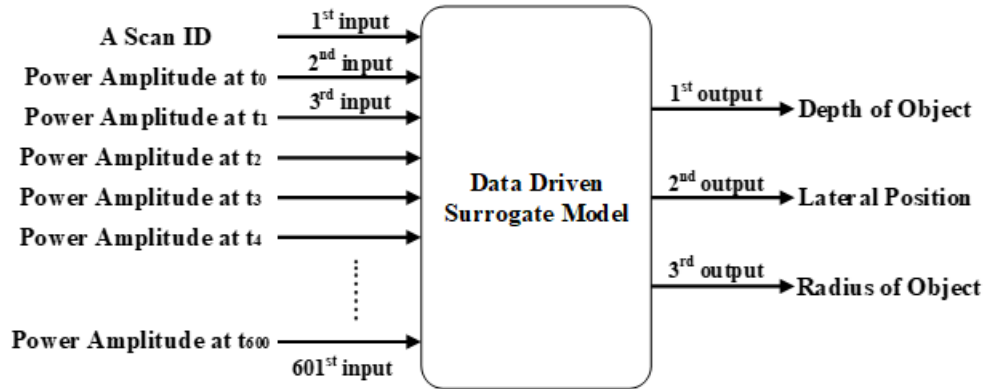


Figure 2.4 Black box representation of the proposed data driven surrogate model for buried object characterization

2.4.2 Generation of Noisy Data Sets for Supplementary Evaluation

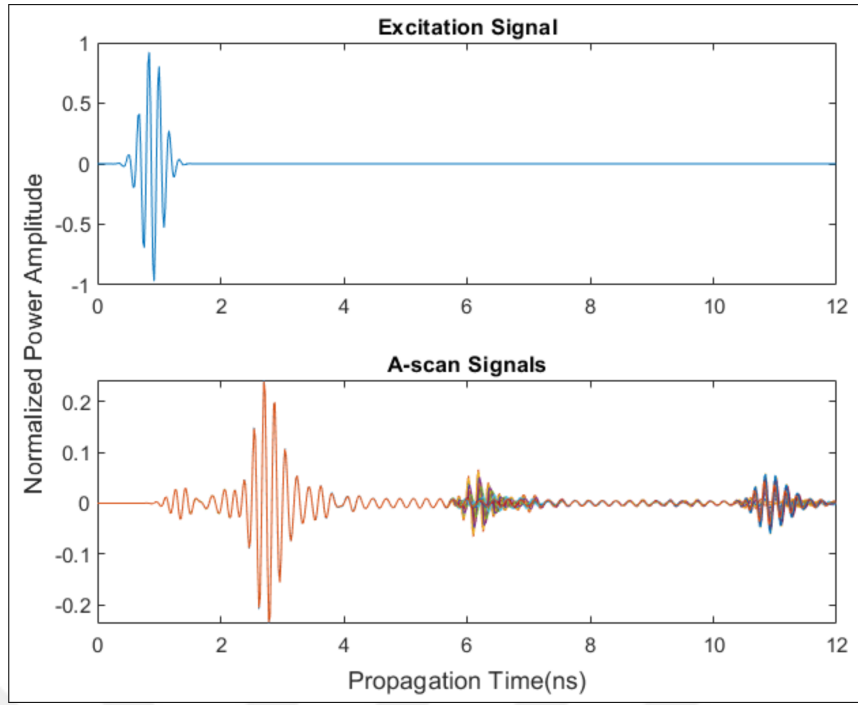
In this study, for the purpose of further verification of the proposed surrogate model, new data sets are generated by random noise addition [14, 17, 25, 32–36] to the generated raw A-scans. The literature offers different approaches to noise incorporation and for different purposes such as data augmentation [14, 16, 24, 32], being closer to realistic scenarios [16, 17, 24, 32, 37] and obtaining further verification to test the sensitivity and stability of the considered models [25, 33–36].

The cases studied in [14, 17] are arranged to bring the models closer to the real-time applications, specifically by considering noisy data sets. A noise with the

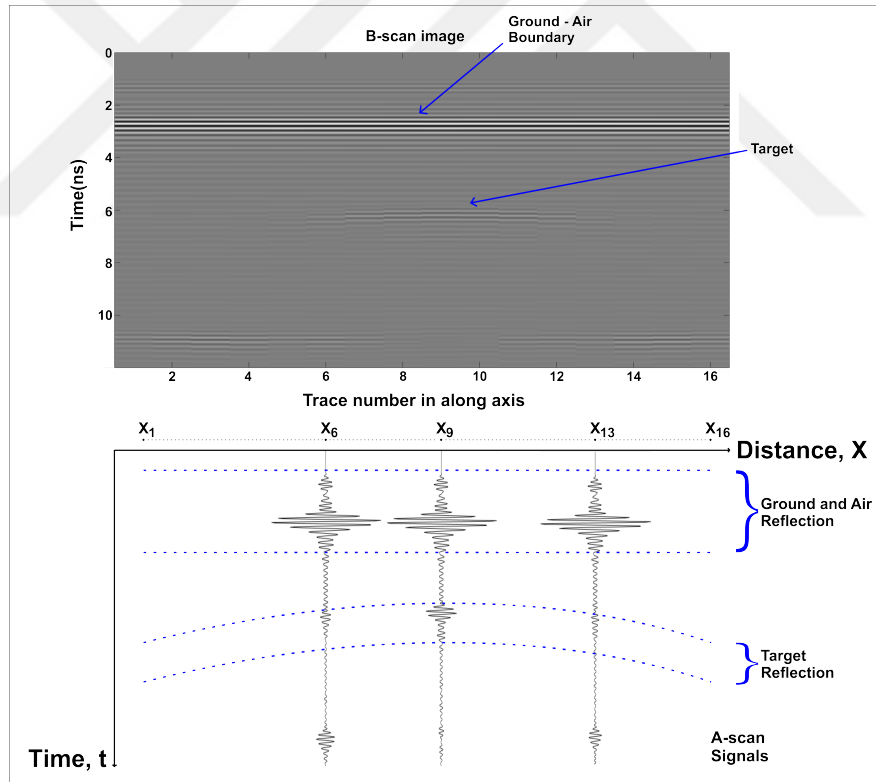
signal-to-noise ratio of 10 dB [17] is added to B-scan image generated by gprMax simulation tool [22, 23]. In [14], random noise is added to B-scans generated by gprMax tool. Another technique is used as integration of the real background reflections with B-scans in order to generalize on realistic scenarios [24] as well as for data augmentation purpose. Application of a similar technique is explained in the study of buried target detection with deep learning [16] by using real background without targets and removed air-soil boundary reflections. Another approach is to replace randomly chosen pixels (typically, from 0.3 % to 25 % of the overall number of pixels) with white and black pixels to obtain noisy data and for further verification of their model in realistic scenarios by using noisy data [37]. This methodology is applied to GPR B-scans generated by gprMax tool and in predicting the object size using a CNN model and extracted hyperbolic patterns [37]. This method is suitable for analysis using B-scan image or 2D data processing techniques, thereby approaching real environmental conditions because reflected waves in concatenated form (B-scan) in different amounts are negatively affected and distorted depending on the percentage of pixel altered.

The particular problem considered in this work is solved by using the A-scan analysis, 1D time-varying amplitude signals (A-scans), and A-scan ID combination, so the methodology of randomly white and black pixel adding is not applicable to the proposed surrogate modeling approach. Hence, for further verification of the proposed surrogate model, supplementary noisy data sets were created with different SNR values (20 dB and 30 dB) by adding white Gaussian noise [25] to emulate conditions that are closer to the real time or on site applications. In Fig. 2.6, noisy A-scans are demonstrated with two different scenarios and different SNR value of 20 dB and 30 dB.

It should be emphasized that the approach to noise incorporation employed in this work is commonly used in the literature. For example, adding Gaussian noise is proposed for an automatic recognition and localization of pipelines [36]. A deep convolutional neural network model is developed to coastal hazard mitigation [32]. Therein, random Gaussian noise is inserted to replicate field scenarios [32]. Some subjects on noise suppression and denoising in addition to the mentioned issues follow the mentioned approaches and utilize Gaussian noise addition [33–36]. In GPR systems, interior or system noise lead to interferences on the reflected signals; it is defined as similar to white Gaussian noise [33]. The K-singular value decomposition (K-SVD) dictionary learning method for the denoising of GPR signals [33] has been introduced, which can effectively suppress Gaussian noise. Another study proposes ensemble empirical mode decomposition (EEMD) method for noise suppression on original GPR data with added white Gaussian noise [34]. The method has been



(a)



(b)

Figure 2.5 (a) Input signal and a sample of raw A-scan signals for water content of 2.8 %, depth 250mm, lateral position 210mm, and radius 40mm (b) B-scan image construction sample for the same scenario

shown successful for denoising [34] synthetic noisy and practical GPR data. Also, different amplitudes of white Gaussian noise are added to generated signals to test the sensitivity to noise and stability of the model [25–35].

To facilitate further research in this area, the data sets (with and without noise) used in this work have been shared in the IEEE data port [38]. To briefly describe the data sets, the training and the testing sets are data matrices of the size of 5040×605 and 1008×605 . In the data sets, 5040 (obtained from 315 different scenarios), and 1008 (obtained from 63 different scenarios) are the sample sizes of training and test data sets, respectively. The features in the 1st and 601st columns are the input of the model while the depth, lateral position, radius of the object, and water content of the soil are present in between 602nd-605th features. As mentioned in Section 2.3, although the water content of the soil is a variable, this feature is not presented to the model since such an approach might not be feasible in practice due to the inhomogeneities in the examined area. In [38], three data set pairs for studied cases in this work are presented with the mentioned data set as explained above: (I) data without any noise, (II) data with 20 dB SNR, and (III) data with 30 dB SNR rate.

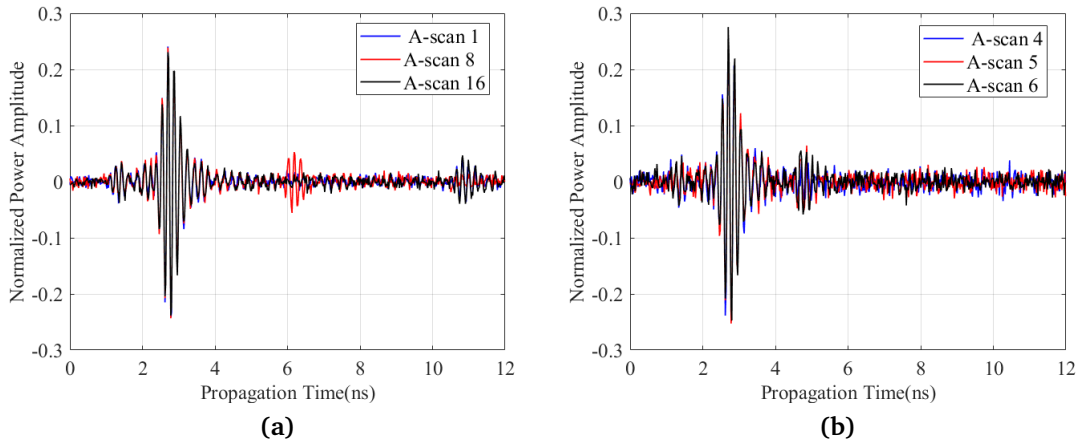


Figure 2.6 Samples of noisy A-scan signals obtained for SNR value of 30dB and 20dB. (a) A-scan samples of SNR 30dB for a scenario of water content of 2.8 %, depth 250mm, lateral position 210mm, and radius 40mm. (b) A-scan samples of SNR 20dB for a scenario of water content of 5.5 %, depth 150mm, lateral position 140mm, and radius 30mm

2.5 Data-Driven Surrogate Models

This section outlines the benchmark techniques utilized in the context of model-based characterization of buried objects, as well as introduces a novel deep-learning-based approach, time-frequency regression model (TFRM). MATLAB has been used as coding

environment for all analysis, training, and testing of the benchmark methods and the proposed TFRM framework.

2.5.1 State-of-the-Art of Surrogate-Based Buried Object Characterization

Surrogate-based characterization of buried objects has been a subject of extensive research over the last years. Some of popular methods utilized in this context include SVRM [39], MLP [3, 7, 9], GP regression [20] and CNN [12, 13, 26, 37]. These methods are briefly characterized below, and will be used as benchmark methods compared to the modelling approach introduced in Section 2.5.2.

Gaussian process approach is an approximation-based machine learning technique widely used for estimation and classification problems. GP regression is based on generalizing Gaussian probability distributions to functions [20, 40]. In these works, the kernel function for numerical experiments has been selected as “matern3/2” [41], and the prediction method has been defined as block coordinate descent with the block size of 250. Determination of hyper-parameters is an important step in surrogate modelling. Here, Bayesian optimization (an in-built optimization tool in MATLAB) has been used for optimum hyper-parameters determination. For validation, the K-fold technique with $K = 5$ has been used.

Another surrogate modeling technique in the machine learning class is support vector regression machine (SVRM), which has been applied not only to object and material type detection through classification [8, 9, 27] by using support vector machine (SVM), but also to prediction of soil permittivity and depth [39] in regression approach. It belongs to the group of supervised statistical learning methods [8, 9, 27, 39].

Herein, similar to the GP regression model, the hyper-parameters of SVRM are determined via Bayesian optimization to realize nonlinear mapping between the reflected received signals and characteristic parameters of the scatterer. The optimally selected hyper-parameters for SVRM are: Gaussian kernel functions, Box-constraint of 1.72, Kernel scale of 0.1977 and Epsilon 0.0008. MLP mimics biological neural systems in the form of interconnected neurons. This framework has demonstrated high accuracy in surrogate modeling applications [3, 7, 9, 11, 42]. The MLP model utilized here features the following hyper-parameter configuration: two hidden layers with 32 and 16 hidden neurons, respectively; log-sigmoid activation functions, training by the Levenberg-Marquardt algorithm. Similar to GP regression and SVRM, these hyper-parameters are also obtained via Bayesian optimization.

The last benchmark surrogate model is CNN, which is a technique derived from deep learning [6, 12, 13, 16, 26]. It takes its name from the convolutional layer, which is one of the main elements in the network, and has the ability to automatically extract the features owing to the convolution filter in this layer. To implement the CNN model, several blocks are utilized as a convolution layer (filter), a batch normalization layer, a pooling layer, the activation function, and a fully connected (FC) layer [43], all involved in the hidden layer. The structure and the hyper-parameter configuration of the CNN used in this work is as follows: three convolution layers (3×1) followed by the batch normalization layer, activation function as ReLU (Rectified Linear Unit) layer, as well as the pooling layers (2×1) incorporated after the last convolution layer, and a fully connected layer with three neurons to model the requested outputs. The CNN model is trained using a version of a back propagation algorithm (referred to as the “Adam” optimizer) and a batch size of 256. The learning rate has been set to 10^{-3} until maximum epoch number reached to 300. Other user-defined parameters, such as the size and the number of the convolution filter (32, 64, 128) and pooling layer are set up as recommended in the literature in the context of buried object detection and characterization [13]. At the same time, it should be mentioned that—in the cited works—the input data is two-dimensional, therefore, the filters of the CNN layers are of the corresponding dimensionality, whereas the filters of the CNN model used in this work are one-dimensional.

All the algorithms considered in this work, including CNN and the proposed technique use the same data sets (1D) that include time-varying reflected normalized power amplitude and A-scan ID (1×1) according to scanning aperture. As mentioned in Section 2.4.1, the data set used for the proposed surrogate modeling is a combination of 1D time-varying amplitude signals (A-scans, 600×1) and A-scan ID (601×1). The first input gives the A-scan ID varying between 1 and 16 according to the along axis. It should be reiterated that the proposed TFRM is a customized deep-learning-based framework that internally converts the given input data as time versus frequency spectrogram, which is the specific form of A-scan and A-scan ID combination.

2.5.2 Proposed Time-Frequency Regression Model Technique for Surrogate Modeling Buried Object

This section introduces the proposed deep-learning-based time-frequency regression model (TFRM) for buried object characterization. The Short Time Fourier Transform (STFT) of 1D signals offer a joint distribution that enables both time and frequency analysis [44]. In addition, the STFT images (2D data) have been used to classify sound signals in many CNN-based studies [45]. The proposed time-frequency regression

model (TFRM) has been inspired by the aforementioned developments. Here, it should be mentioned that deep learning algorithms have shown great success in problems featuring large data sets (e.g., in big data applications). It is worth mentioning that applicability and success rate of deep learning is not strictly bounded to cases featuring large numbers of training samples [46]. Deep learning algorithms can be used to transform the input data to a higher-dimensionality space using large number of neurons and layers, and then distilling the transformed data into lower dimensions to better handle the interrelations between the problems input and outputs. Thus even with small amount of training samples it is possible to use deep learning algorithms to create a globally accurate surrogate model [47].

As mentioned earlier, the proposed methodology utilizes GPR data as time-varying signals. In other words, A-scans and A-scan analysis is carried out. The A-scan analysis including data obtained from B-scans is designed by adding A-scan ID (1×1) according to the lateral position at the scanning aperture (along axis) to each of the A-scans of all of the B-scans. Here, the A-scan ID is an integer between 1 and 16. In this way, 2D data (B-scan) is reduced to 1D data combination of reflected field amplitude versus time, and A-scan ID (601×1), so that A-scan data [38] with a length of 601 is used in this study. It should be emphasized that the proposed TFRM (time-frequency regression model) is a customized deep-learning-based framework transforming internally the given input data as time versus frequency spectrogram (2D time-frequency spectrogram), which is the specific form of A-scan and A-scan ID combination. It should be mentioned that the literature offers a similar approach [13], oriented towards detection of buried explosive object and discrimination target or nontarget alarms. Therein, 3D GPR volume data is used as an input by converting 2D data.

Firstly, the signals are zero-averaged. Then, the blocks of the length 32 are extracted over sixteen overlapping intervals to extract the STFT image of the signals. Each block is multiplied by the Kaiser window [48] to prevent the spectral leakage. Subsequently, Fast Fourier Transforms (FFT) of the length of 64 are taken. Since the FFT is symmetrical for real signals, the one-sided portion of the image is taken. Thus, a spectrogram of a dimension 33×36 is obtained. The magnitude spectrum is calculated to process the spectrogram as an image. With the conversion of complex numbers to real numbers by taking their magnitude, the STFT magnitude spectrum images of A-scan signals are formed. Next, these images are reduced to dimension 32×32 to be an exact multiple of two (cf. Fig. 2.7). Before the model is trained, the A-scans are zero-averaged. The TFRM model consists of five main blocks as shown in Fig. 2.8. The first three blocks are used to extract the features from the $32 \times 32 \times 1$ STFT image. After extracting features of the STFT image in the convolution layers,

data processing is followed with the batch normalization operations and Leaky ReLU layer with the scale parameter of 0.1 (a scalar multiplier of negative inputs) as an activation function. Owing to the global average pooling layer [49] used in the fourth block, a $1 \times 1 \times 256$ feature vector is obtained. In the last, fifth block, the feature vector is processed over 512 neurons in the fully connected layer, and converted to characteristic parameters in terms of the location (depth, lateral position), and the radius of the object.

The Mean Absolute Error (MAE) and the Relative Mean Error (RME) has been used as performance metrics of the surrogate models. The models are defined as

$$MAE = \frac{1}{N} \times \sum_{i=1}^N |P_i - T_i| \quad (2.3)$$

$$RME = \frac{1}{N} \times \sum_{i=1}^N \frac{|P_i - T_i|}{|T_i|} \quad (2.4)$$

where N is the total number of samples, whereas T_i and P_i are the target and model-predicted values, respectively, for the i th sample.

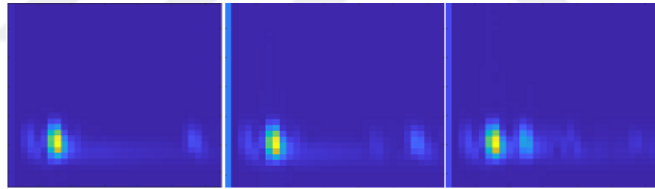


Figure 2.7 The obtained 32x32 time-frequency images for different signals

Table 2.3 shows the average errors of all considered surrogate models and their corresponding average training times, averaged over ten independent runs. In order to clearly demonstrate the performance of the proposed TFRM, it is compared with the latest developed and most commonly used regression methods reported in the literature (GP regression, SVRM, MLP, CNN). To ensure a fair comparison between TFRM and mentioned counterpart algorithms, the hyper parameter configuration of these models are taken as standard values in literature or based on the counterpart works given configurations [13]. It can be observed that the accuracy of the proposed TFRM framework is significantly better (given the training datasets, which are identical for all models) than all benchmark surrogates. In particular, the MAE of TFRM is twice as low as for the second-best model (here, CNN). Furthermore, the training time of TFRM is considerably lower than for all benchmark methods.

Table 2.4 shows the modelling errors for individual characteristic parameters for the

proposed and the benchmark models. It can be observed that, in a qualitative sense, the results are consistent with those presented in Table 2.4. As it can be seen from the error metrics of the benchmark models and the proposed model, the performance of parametric estimation of geophysical parameters of buried object in terms of depth, lateral position and radius independent of each other with raw A-scan signal analysis is satisfactory, and it is achieved in a computationally efficient manner even in the case of using sparse training samples. Also, Table 2.5 and Figure 2.10 demonstrate the target and predicted parameters of selected test samples for the second-best benchmark model and the proposed model.

The TFRM framework uses the Adam optimizer, which is a version of the back propagation algorithm for training the framework. Here, it is used with the batch size of 256 and the maximum epoch number of 300, also data is shuffled in every epoch. In Figure 2.9 training progress has been demonstrated with train loss values versus the iteration number.

Table 2.3 Accuracy and the training time of the proposed and the benchmark models, averaged with standard deviation over ten independent runs

Model	MAE [mm]	RME [%]	Training Time [Minutes]	Testing for a Single Input Data [ms]
GP regression	37.6±1.2	31.7±0.7	8	4.9
SVRM	41.1±1.1	38.0±1.4	44	23
MLP	37.4±2.2	31.2±1.1	150	5
CNN	24.7±1.0	27.5±0.6	70	9
TFRM [this work]	12.1±0.1	16.5±0.5	5	6

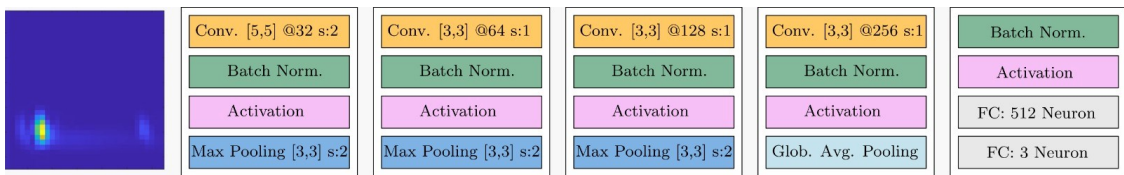


Figure 2.8 Architecture of the proposed TFRM model

2.6 B-scan Evaluation

In the proposed approach, the buried objects characterization is obtained using the 1D data of the normalized power amplitude of the reflected signals. The latter are obtained for different positions of the scanning aperture, by means of the proposed TFRM framework. For additional validation, and to better demonstrate the performance of the proposed approach as compared to the benchmark methods, a B-scan evaluation is studied in this section.

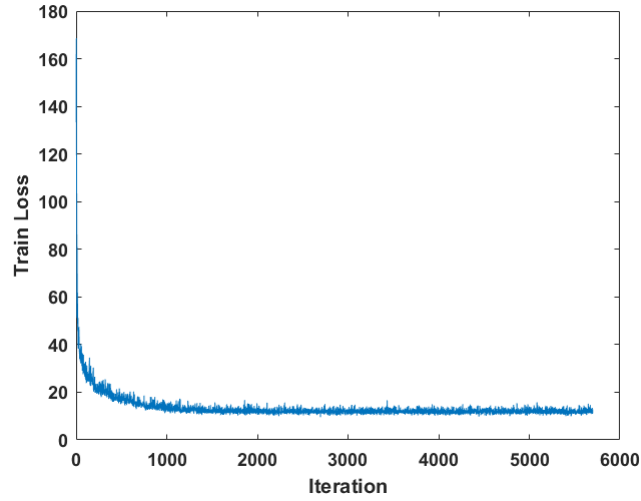


Figure 2.9 The training loss of the proposed TFRM framework versus the iteration number during the all epochs (epoch number equals to 300)

The B-scan images are 2D datasets that can be generated by combining the A-scan data obtained from different positions [1–12]. In the GPR setup defined in Section 2.3, a B-scan data can be obtained using sixteen A-scan datasets for each scenario. Again, it should be mentioned that by increasing the total number of A-scans, it is possible to increase the accuracy of the model; however, this will be at the expense of decreasing the computation efficiency. The performance evaluation of the best benchmark method, CNN, and the proposed approach, TFRM, have been presented in Table 2.5, and Figure 2.10.

Table 2.4 Breakdown of the best prediction performance of characteristic parameters

Model	Characteristic	MAE [mm]	RME [%]	Average MAE [mm]	Average RME [%]
GP regression	Lateral Position	28.8	14.3	35.6	30.8
	Radius	8.8	45.1		
	Depth	69.3	33.0		
SVRM	Lateral Position	37.6	18.8	39.0	36.1
	Radius	11.7	58.2		
	Depth	67.7	31.3		
MLP	Lateral Position	35.5	18.0	34.3	29.9
	Radius	8.6	44.3		
	Depth	59.0	27.4		
CNN	Lateral Position	38.1	19.2	23.0	26.5
	Radius	10.7	51.5		
	Depth	20.1	8.9		
TFRM [this work]	Lateral Position	17.7	8.7	11.9	15.9
	Radius	7.7	34.5		
	Depth	10.4	4.7		

These results are the average values for all characteristic parameters (depth, lateral position, radius) obtained from sixteen different A-scans, compared with the actual object characteristics. As it can be seen, the performance of the proposed TFRM framework (average MAE of 11.95) is clearly superior in terms of object characterization over the best benchmark technique, CNN (average MAE of 23.01).

It should be noted that TFRM can almost pinpoint the target object center position, whereas CNN fails with this respect.

As mentioned in Section 2.5.2, for the supplementary evaluation of performance of the proposed model by using noise addition to raw A-scans, new data sets were generated by adding a random noise [14, 17, 25, 32–36] to the A-scan signals. Thus, the environmental and internal noise of the GPR system and its effect was analyzed. In particular, the supplementary noisy data sets were created with different SNR values (20 dB and 30 dB) by adding the white Gaussian noise [25]. These noisy data sets have been analyzed without any pre-processing to estimate characteristic parameters under scenarios that are closer the real time or on site applications. For this investigation, the second best successful surrogate model, CNN, was used to compare the performance of the proposed framework and the results have been presented in the following table. Also, performance evaluation of the best benchmark model CNN, and the proposed approach TFRM for the noisy data sets have been demonstrated in Table 2.7 and Figure 2.11.

Here, it must be emphasized that all the data samples in training and test data sets, 315 scenarios with linear sampling and 63 scenarios with Latin Hypercube Sampling (LHS) method, are completely different from each other. This can be observed in Fig. 2.3 and the obtained results. In the case of high similarity between the training and test samples or high amount of training samples compared to test samples, all surrogate models would exhibit high accuracy (either true performance or over-fitting). In this case, since the same data samples are utilized to construct all models (GP regression, SVRM, MLP, CNN, TFRM) it is expected that all methods present similar high accuracy in case of training and testing sample similarity. However, the results presented in Tables 2.4-2.7 and Figs. 2.10, 2.11 suggest the opposite.

Furthermore, with respect to results in Tables 2.4 and 2.6, in case of increasing the complexity of the data set such as adding noise, the performance of TFRM is still superior to the nearest benchmark method (CNN) where the average MAE of TFRM increase from 11.9 to 27.5 for SNR value of 30 dB, the CNN average MAE is increased from 23 to 38.3. Thus, even though increasing the complexity of the data set either via changing the training test ratio or adding noise would definitely have an impact

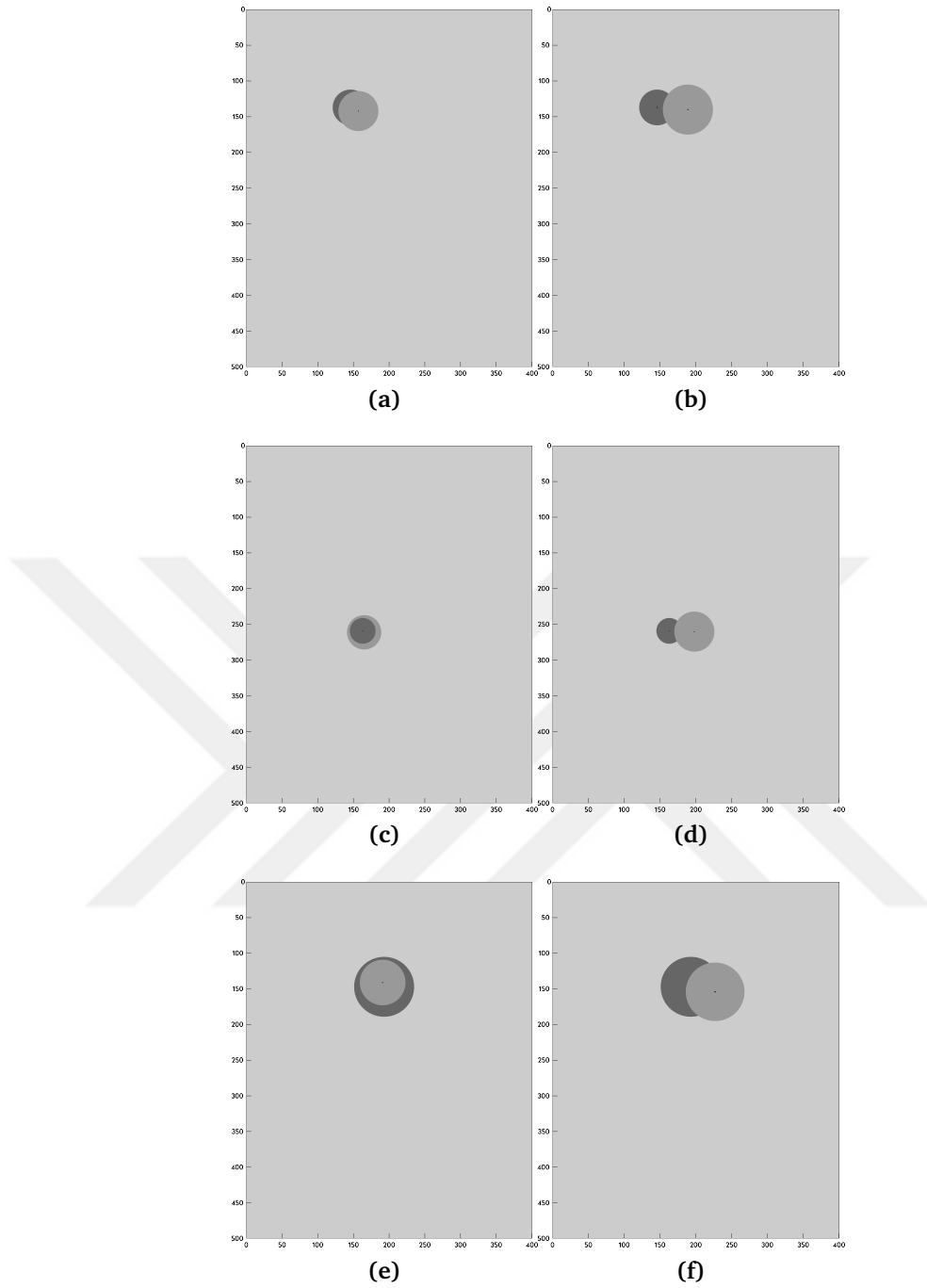


Figure 2.10 Inversion results of surrogate modeling with TFRM and 1-D CNN for B-scan evaluation of raw time signals, the image where the dark gray shows target and light gray predicted in the geometry of subsurface, (a) ($D = 137$ mm, $P = 146$ mm, $R = 25$ mm) estimated by TFRM, (b) ($D = 137$ mm, $P = 146$ mm, $R = 25$ mm) estimated by CNN, (c) ($D = 259$ mm, $P = 163$ mm, $R = 18$ mm) estimated by TFRM, (d) ($D = 259$ mm, $P = 163$ mm, $R = 18$ mm) estimated by CNN (e) ($D = 147$ mm, $P = 193$ mm, $R = 42$ mm) estimated by TFRM, (f) ($D = 147$ mm, $P = 193$ mm, $R = 42$ mm) estimated by CNN

on the overall performance of all models that can be seen from the presented results, and under such case still the proposed TFRM is superior to the benchmark methods.

In the proposed model, the characteristic parameters of the object are estimated without any information of the background media such as dielectric permittivity or water content which is in a proportionality relation with the dielectric permittivity and conductivity of the underground according to the Debye model. This approach is followed to account for complexity of the generated scenarios with regard to the relative dielectric permittivity in different subsurface medium being dependent on the water content. As defined in the study of reconstruction of permittivity images with deep convolutional network [26], real dielectric features of the underground medium are complex and it is difficult to obtain. For this reason and because of the relationship between the traveling time of signals (the depth) and the subsurface permittivity, the error rates in the estimation of geophysical parameters of the object can be negatively affected.

Table 2.5 TFRM- and CNN-based prediction of the characteristic parameters in comparison with their true values for the selected test scenarios

Model	<i>D</i>	<i>P</i>	<i>R</i>	Error (<i>D, P, R</i>) [mm]
True Value	259	163	18	-
TFRM [this work]	261	165	24	-2, -2, -6
CNN	260	198	28	-1, -35, -10
True Value	137	146	25	-
TFRM [this work]	142	157	28	-5, -11, -3
CNN	140	189	35	-3, -43, -10
True Value	376	155	30	-
TFRM [this work]	376	156	28	0, -1, -2
CNN	394	195	36	-18, -40, -6
True Value	204	241	27	-
TFRM [this work]	209	235	27	-5, 6, 0
CNN	210	227	36	-6, 14, -9

By adding water content parameters to input data, an additional analysis has been conducted to investigate the influence of knowledge about the background media on the proposed model. The resulting average MAE of 11.1 [mm] and the average RME of %15.30 has been obtained.

The characteristic parameter prediction is as follows: the radius predicted with MAE of 7.3 [mm], RME of %33.2; the depth predicted with MAE of 9.2 [mm], RME of %4.3, and the lateral position predicted with MAE of 16.59 [mm], RME of %8.3. These numbers indicate that the incorporation of the additional knowledge has a positive effect on the error metrics.

Table 2.6 Prediction performance of considered characteristic parameters for noisy data sets

Model	Characteristic	MAE [mm]	RME [%]	Average MAE mm]
CNN [30dB]	Lateral Position	43.8	21.3	38.3
	Radius	10.2	46.7	
	Depth	60.9	26.4	
TFRM [30dB]	Lateral Position	27.6	13.8	27.5
	Radius	8.5	43.9	
	Depth	46.3	21.2	
CNN [20dB]	Lateral Position	43.1	20.9	41.2
	Radius	10.3	47.6	
	Depth	70.1	30.3	
TFRM [20dB]	Lateral Position	29.6	15.0	31.4
	Radius	8.6	44.1	
	Depth	55.9	26.3	

Table 2.7 TFRM- and CNN-based prediction of the characteristic parameters in comparison with their true values for the selected test scenarios for noisy data sets

Model	<i>D</i>	<i>P</i>	<i>R</i>	Error (<i>D, P, R</i>) [mm]
True Value	259	163	18	-
TFRM [30 dB]	264	186	28	-5, -23, -10
CNN [30 dB]	226	213	29	33,-50,-11
True Value	137	146	25	-
TFRM [20 dB]	204	178	30	-67, -32, -5
CNN [20 dB]	217	207	31	-80, -61, -6
True Value	376	155	30	-
TFRM [30 dB]	292	195	30	84, -40, 0
CNN [30 dB]	257	193	28	-119, -38, 2
True Value	204	241	27	-
TFRM [20 dB]	214	231	29	-10, -10, -2
CNN [20 dB]	216	197	26	-12, 44, 1

2.7 Validation of Data-Driven Surrogate Modeling Using Measurement Data

In this chapter, the performance of the proposed surrogate modeling approach was also validated using measurement data. 1D A-scans have been collected through the measurements in a “sand pool” environment. Herein, the purpose is to demonstrate that the proposed approach is also applicable when using physical measurements as a source of data. The experimental samples are obtained in the laboratory at Yıldız Technical University. During the process, raw B-scan data corresponding to various scenarios are generated by the impulse ground penetrating radar system, which is utilized in various subsurface imaging operations [50–53]. Figure 2.12 shows

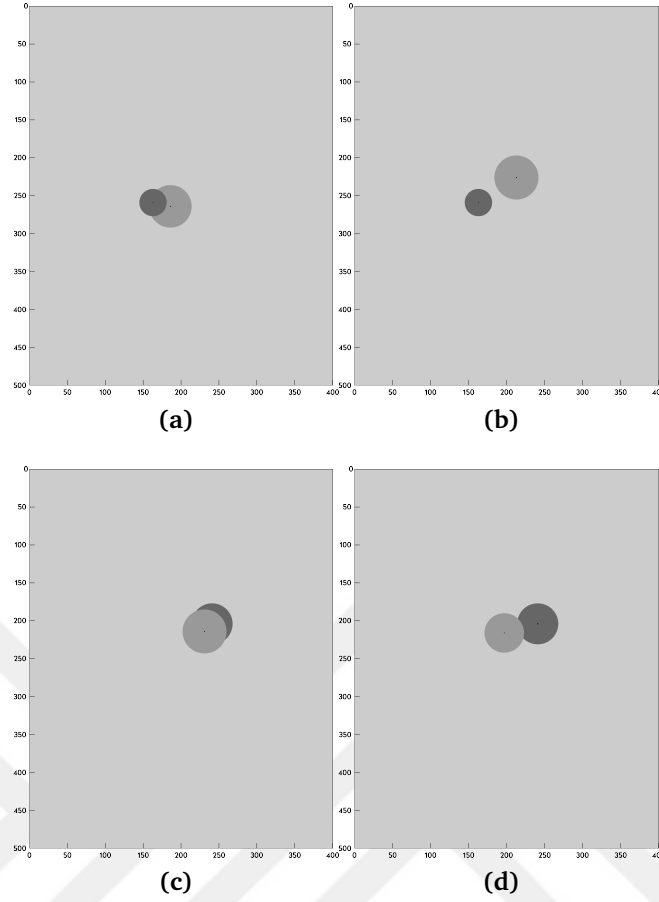


Figure 2.11 Inversion results of surrogate modeling with TFRM and 1-D CNN for B-scan evaluation of noisy raw time signals, the image where the dark gray shows target and light gray predicted in the geometry of subsurface, (a) ($D = 259$ mm, $P = 163$ mm, $R = 18$ mm) estimated by TFRM [30 dB], (b) ($D = 259$ mm, $P = 163$ mm, $R = 18$ mm) estimated by CNN [30 dB], (c) ($D = 204$ mm, $P = 241$ mm, $R = 27$ mm) estimated by TFRM [20 dB], (d) ($D = 204$ mm, $P = 241$ mm, $R = 27$ mm) estimated by CNN [20 dB]

the experimental setup. The measurements are taken in a wooden pool filled with inhomogeneous dry soil consisting of a mixture of small stones and sand. The object is buried in there used as the scanning subsurface domain. It has the dimensions of approximately 1.40 m (width), 0.22 m (depth) and 1.15 m (length).

The experimental setup (GPR, transmitter and receiver antennas) is manually moved above the soil along the scanning path of the approximate length of 1.40 m. Each B-scan is set as 382 (discrete time step) $\times$ 65 (A-scan number). Hence, the input data length is 383, the first input data being the A-scan ID according to lateral position at the scanning axis, and the left part of the input being the amplitude of reflected time signals (382 discrete time steps). Figure 2.13 presents A-scans from a sample test scenario. The buried object which has various radii of 10mm, 15mm, 20mm, 25mm and 30mm is placed at different locations (depth and lateral positions as 550mm,

650mm, 770mm and 900 mm). The data set created using 33 scenarios to train and test of the proposed surrogate model on the measured A-scans.

Two case studies are used to investigate the performance of the proposed surrogate modeling approach for measurement data. The first one (Case 1) is based on K-fold validation [54]. The second one (Case 2) is when the available data is split into training and hold-out data sets.

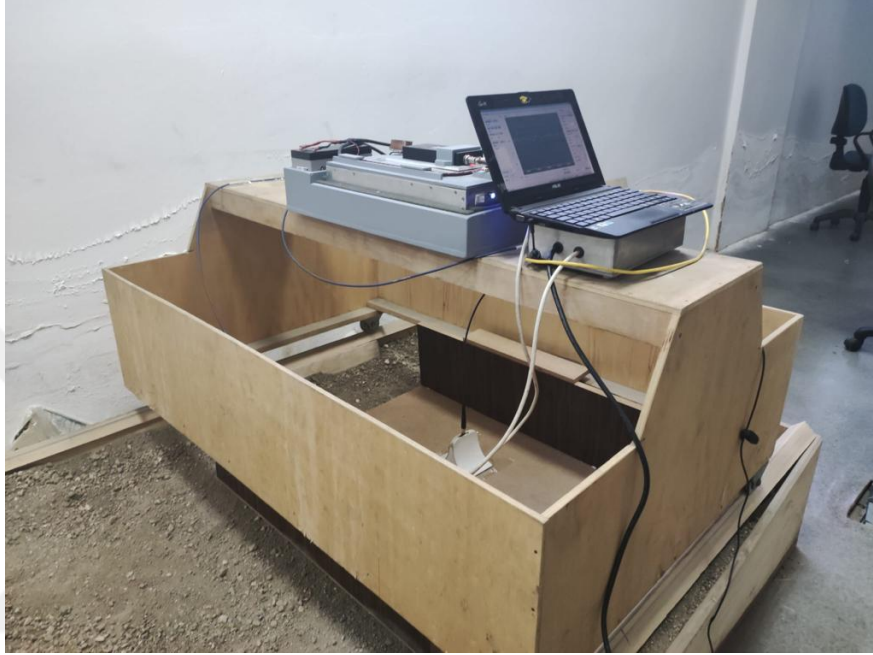


Figure 2.12 The picture of the experimental environment and setup used to generate the measurement data set

The unprocessed raw B-scan images are mapped to 1D data (A-scan and A-scan ID combination, 383×1) for the use in the proposed A-scan analysis technique to simultaneously obtain characteristic parameters of the buried object in terms of its depth, lateral position and radius. For the sake of comparison, the CNN method has been used as a benchmark technique. The results can be found in Tables 2.8 and 2.9, as well as in Fig. 2.14.

These results are the average values for all characteristic parameters (depth, lateral position, radius) obtained from different runs of the models. Figure 2.14 and Table 2.10 demonstrate the predicted characteristic parameters compared with the actual object parameters. As it can be seen, the performance of the proposed TFRM framework (average MAE of 25.9) is clearly better in terms of object characterization than the benchmark (average MAE of 35.0). In particular, TFRM-based prediction of the geophysical parameters is satisfactory as indicated in Table 2.10 and Fig. 2.14, and it is considerably better than that of CNN.

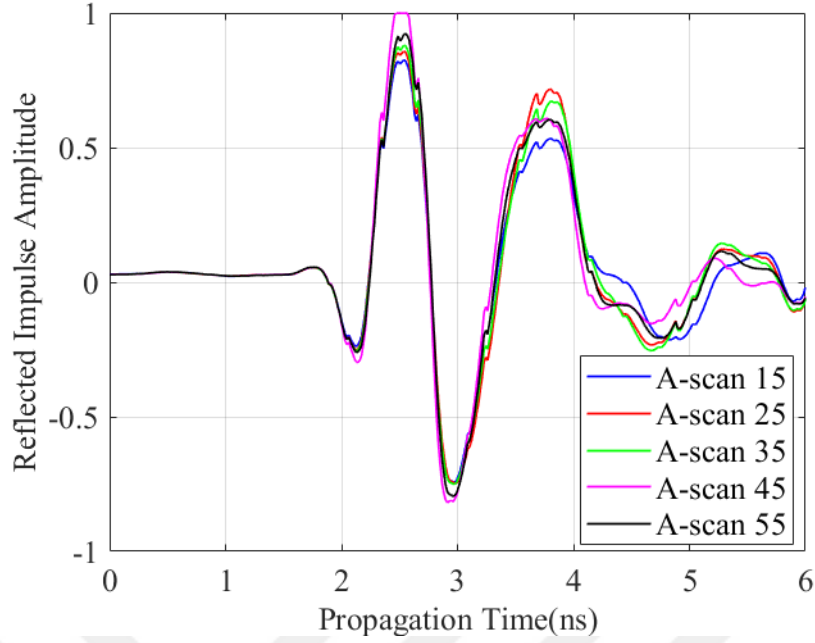


Figure 2.13 Samples of measured A-scan signals obtained from the sand pool by using the experimental set up. A-scan samples belong to a scenario given by the object depth 115mm, lateral position 550mm, and radius 25mm

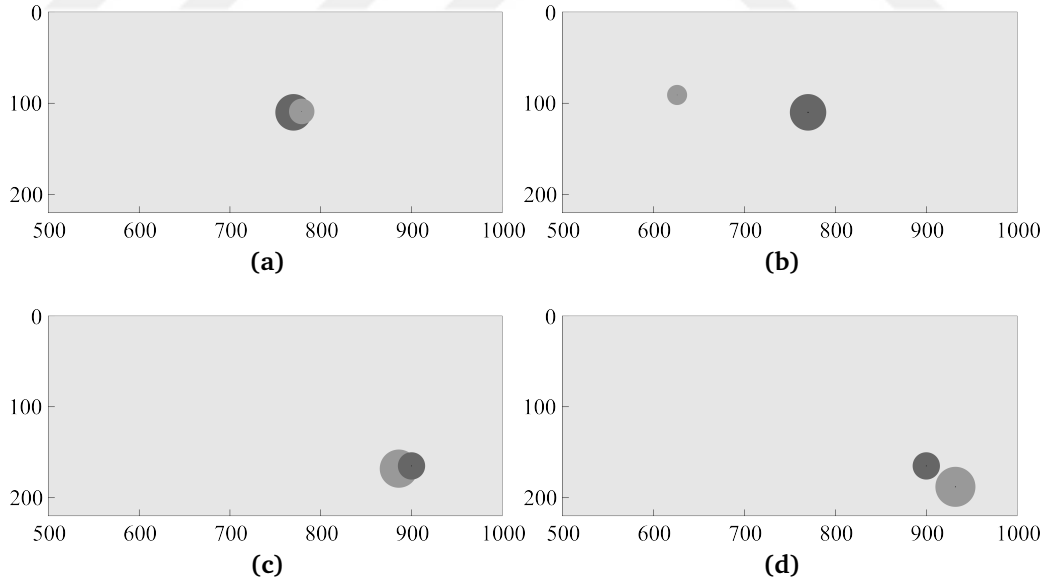


Figure 2.14 Inversion results of surrogate modeling with TFRM and 1-D CNN for B-scan evaluation of hold-out measurement data, the image where the dark gray shows target and light gray predicted in the geometry of subsurface, (a) ($D = 110$ mm, $P = 770$ mm, $R = 20$ mm) estimated by TFRM, (b) ($D = 110$ mm, $P = 770$ mm, $R = 20$ mm) estimated by CNN, (c) ($D = 165$ mm, $P = 900$ mm, $R = 15$ mm) estimated by TFRM, (d) ($D = 165$ mm, $P = 900$ mm, $R = 15$ mm) estimated by CNN

Table 2.8 Accuracy and the training time of the proposed and the benchmark model, CNN. Shown are the averaged error metrics with standard deviation obtained for ten independent runs using the measurement data

Model	MAE [mm]	RME [%]	Training Time [Minutes]	Processing Time for a Single Input Test Data [ms]
CNN (Case 1)	42.0±2.7	30.6±1.1	18	9
CNN (Case 2)	37.6±1.2	31.7±0.7	18	9
TFRM (Case 1) [this work]	34.9±0.7	29.1±1.1	4	6
TFRM (Case 2) [this work]	28.1±1.2	18.7±1.1	4	6

Table 2.9 Breakdown of the prediction performance of the considered characteristic parameters for hold-out data

Model	Characteristic	MAE [mm]	RME [%]	Average MAE [mm]
CNN (Case 2)	Lateral Position	76.8	10.2	35.0
	Radius	7.0	46.5	
	Depth	20.8	14.4	
TFRM (Case 2) [this work]	Lateral Position	63.9	8.8	25.9
	Radius	6.6	43.6	
	Depth	7.1	4.9	

Table 2.10 TFRM- and CNN-based prediction of the characteristic parameters in comparison with their true values for the selected hold-out scenarios for measurement data set

Model	<i>D</i>	<i>P</i>	<i>R</i>	Error (<i>D</i> , <i>P</i> , <i>R</i>) [mm]
True Value	100	650	10	-
TFRM [this work]	104	651	17	-4, -1, -7
CNN	114	653	21	-14, -3, -11
True Value	170	550	20	-
TFRM [this work]	164	551	23	6, -1, -3
CNN	132	547	20	38, 3, 0
True Value	180	650	30	-
TFRM [this work]	170	618	21	10, 32, 9
CNN	138	594	18	42, 56, 12

2.8 Conclusion of Buried Object Characterization with A-scan Analysis and Data-Driven Surrogate Models

Herein, introduced a novel approach to surrogate-assisted characterization of buried objects. The presented approach utilized AI methods, in particular, a deep-learning-based TFRM framework. Its major advantage is computational

efficiency, and the ability of constructing accurate representation of the buried object characteristics using small training datasets. The presented framework has been comprehensively validated using a number of specific cases of perfectly electric conductor (PEC)-object buried in different, dispersive soil media, at various positions. It has also been compared to a number of the latest developed benchmark methods, including GP regression, MLP, SVRM, and CNN, all commonly used for the same purpose. The results indicate competitive performance of the proposed technique with the MAE value of less than 12 (as compared to MAE of 23 for the best benchmark approach, CNN). These results also demonstrate that the TFRM framework can be viewed as efficient and accurate approach to solving GPR-based buried object characterization task under low computational budget. The proposed methodology for estimating characteristic parameters has been validated using noisy data (SNR value of 20 dB and 30 dB) and the measurement data. The error metrics and geometrical representations of characteristic parameters indicate that a novel regression surrogate model, TFRM is superior to the benchmark models, including CNN. The future work will include the extension of the proposed data-driven surrogate modeling approach to characterization of multiple objects and the increased number of characteristic parameters of the buried objects.

BURIED OBJECT CHARACTERIZATION VIA REGRESSION-ENABLED HYPERBOLIC SIGNATURE EXTRACTION AND DATA-DRIVEN SURROGATE MODELS

3.1 Introduction to Buried Object Characterization via Regression-Enabled Hyperbolic Signature Extraction & Problem Formulation

The work described in this chapter proposes a novel methodology for buried object characterization in a computationally-efficient approach via use of data-driven surrogate modelling. The ground reflections are removed, and background subtraction operations are executed so that B-scan image processing can be carried out. The surrogate is constructed at the level of hyperbolic patterns extracted using linear regression [55, 56]. The latter results in dimensionality reduction of the parameter space [57] and enables efficient analysis of GPR data using neural network models. Further, 2D B-scan data is reduced to 1D data consisting of electric field amplitude values along the scanning axis, so that an advantage is obtained with regard to computational cost of the proposed surrogate modeling approach. The estimated hyperbolic patterns contain information about the geophysical properties of the buried object, including its lateral position, depth, and radius. The underlying surrogate modelling technique involves deep-learning-based modified multilayer perceptron (M2LP) framework. Its architecture, in terms of deep-learning-based layers, is similar to a regression model employed to represent scattering parameters of a capacitively-fed antenna [58]. There are two fundamental contributions of this work. The first one is the employment of linear regression techniques for extracting hyperbolic signatures from the B-scan data, which are associated with the geophysical parameters of the buried object, and allow for dimensionality reduction of the dataset being handled in the process. The second is a novel deep-learning-based M2LP framework, which enables significant improvement of the computational efficiency of surrogate model rendition, thereby expediting the object characterization process.

In this section, the buried object characterization task is also formulated. The problem at hand is to estimate characteristic parameters of a buried cylindrical object, specifically its depth D , lateral position P and radius R . The meaning of these parameters has been explained in Fig. 3.1. The object itself is assumed to be made of perfect electrical conductor (PEC). The estimation of the object parameters is realized here using fast surrogate model established using pre-processed data obtained from the GPR model to be discussed later. Data pre-processing allows for a considerable reduction of the dataset complexity, including a reduction of its dimensionality, thereby facilitating the modeling process. This chapter is carried out using a dedicated M2LP framework elaborated on below.

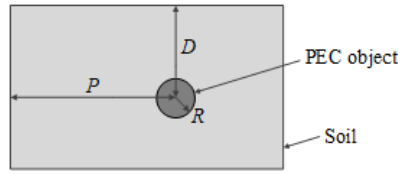


Figure 3.1 Buried object characterization problem: graphical illustration and definition of terms. Variables R , D , and P stand for the object radius, its depth, and lateral positions, respectively, all to be identified in the characterization process

The remaining part of the chapter is arranged as follows. The next sections provide a brief explanation of the GPR model, and the arrangement of the training samples. Another section elaborates on a regression-based hyperbolic pattern extraction from the B-scans, as well as deep-learning-based modified MLP (M2LP) framework developed to perform characterization of depth, lateral position, and radius of the detected object. The following part extends our approach to handle noisy datasets, whereas the follow-up section also extends it to measurement data sets. The last section concludes the chapter.

3.2 Configuration of GPR Model

In this work, the data used for object identification is obtained from the computational model representing the ground penetrating radar (GPR) together with the associated environment (a soil section containing the object buried therein). The GPR model is evaluated using the electromagnetic simulation software gprMax that involved Finite Difference Time Domain (FDTD) solver [22, 23]. The geometrical configuration of the model has been presented in Fig. 3.2. As mentioned earlier, the object is represented as perfect electric conductor (PEC) in the form of a wire, pipe, or rebar. The travelling time t_d of the wave transmitted by the antenna can be computed as

$$t_d = \frac{\sqrt{\epsilon_r} 2d}{c} \quad (3.1)$$

where d is the object depth, ϵ_r is the relative permittivity of the subsurface, and c is the speed of light in the free space. It is similar with the equation (2.1). Note that the wave propagation time is monotonically dependent on the depth but also the subsurface permittivity.

The GPR model is configured as follows:

- The subsurface domain is defined as dry soil of relative permittivity 3 and conductivity 0.001 S/m;
- Subsurface domain dimensions are 0.4 m (lateral) and 0.6 m (vertical);
- The cubic cell size for the spatial discretization of the simulation environment is 1 mm;
- The boundary conditions are defined to be Perfectly Matched Layers (PML) with the thickness of ten cells (at all domain faces);
- The transmitter and receiver antennas (marked at T and R in Fig. 3.2, respectively), are placed 75 mm (almost 0.66λ [8, 20, 21], λ being a guided wavelength) apart from each other and close to the ground surface (the distance is set to 2 mm [21].);
- The radius R of the buried object is assumed to be within the range $10\text{mm} \leq R \leq 40\text{mm}$.

The GPR model is used to generate so-called A-scans [1–3, 6, 10], which are waveforms of the electric field strength recorded by the receiver antenna over the 7.5 ns window. The antennas are moved for each trace (A-scan) at the along axis (scanning path) as shown in Fig. 3.2. The source signal $w(t)$ is set to be a normalized first-order derivative of a Gaussian waveform with a center frequency $f_c = 1.5$ GHz. We have (the unit of time t are seconds)

$$w(t) = -2\pi f_c \sqrt{e} \left(t - \frac{1}{f_c}\right) e^{-2\pi^2 f_c^2 \left(t - \frac{1}{f_c}\right)^2} \quad (3.2)$$

Fig. 3.3(a) shows exemplary A-scans obtained for a specific scenario. The process of constructing the B-scans from the A-scans has been illustrated in Fig. 3.3(b). In

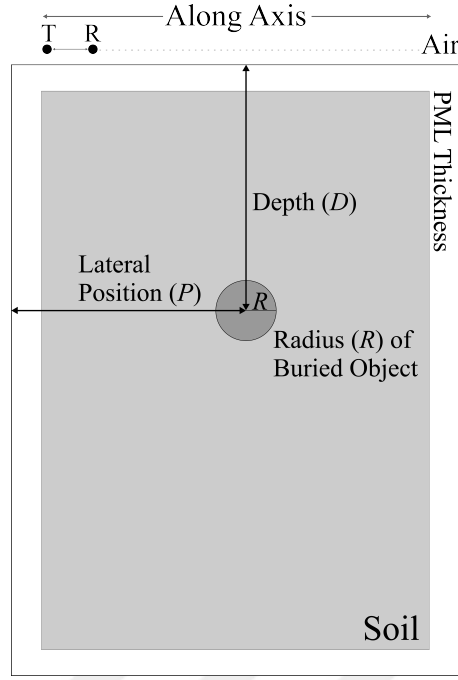
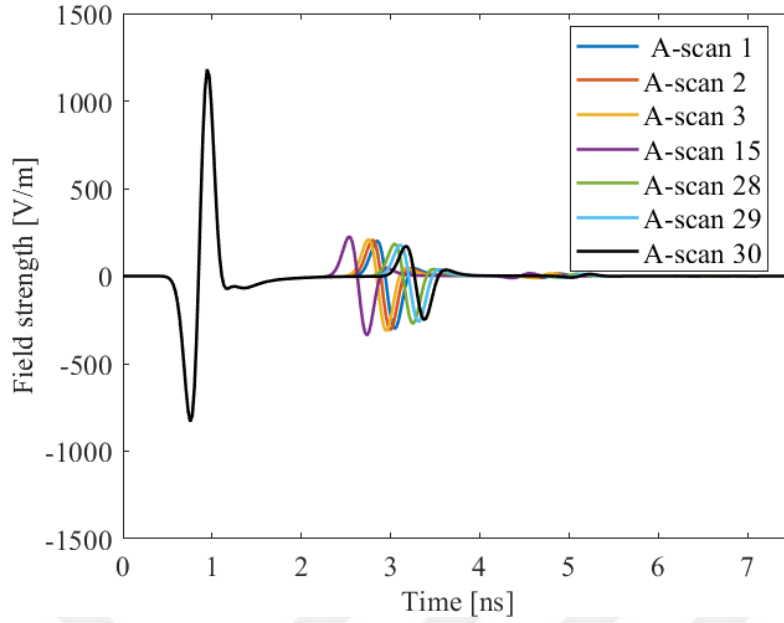


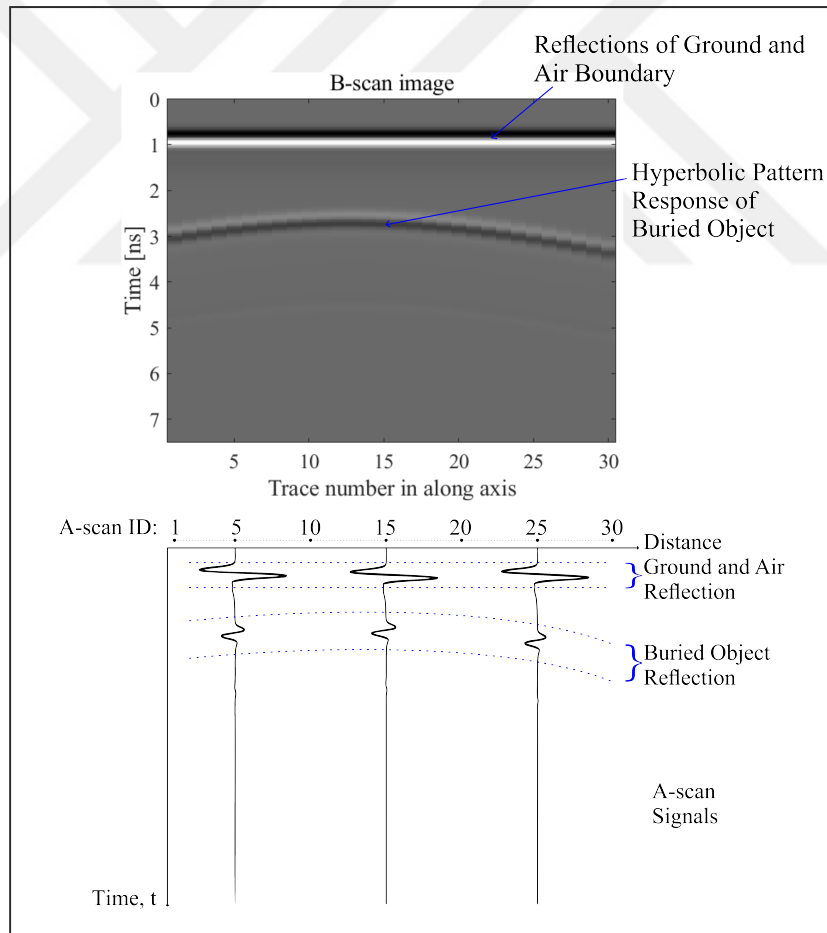
Figure 3.2 Configuration of the GPR model for generating training and testing data used by the proposed surrogate-assisted buried object characterization framework. T and R stand for the transmitter and receiver antenna, respectively. As mentioned earlier, the parameters to be estimated are the lateral position P , depth D , and object radius R

addition, -3 dB level interval of the used waveform is approximately 1.1 to 2.7 GHz, and the approximate frequency bandwidth is 1.6 GHz, which is provided at the same time via the inverse of the pulse width gave bandwidth for the impulse radar system [1, 2]. According to time window ($T = 7.5$ ns) of the one-point scanning process, frequency bin spacing (Δf) at the bandwidth (f_{BW}) is approximately as 133.4 MHz ($1/T$), the frequency step number, N_s is calculated 120 by using $f_{BW}/\Delta f$. The range resolution [1, 2, 10] is also found as 54 mm by using the Fourier theory, which explains the relation between range resolution with velocity of the waveform in the soil medium as directly proportional to the range resolution and with frequency bandwidth as inversely proportional to the range resolution [10].

The maximum penetration depth [1, 2, 10] for this configuration of the model is calculated as 0.649 m by using dielectric features of the subsurface, the value of frequency bin spacing, and the number of frequency steps. The maximum depth for object detection is related to the features of the excitation waveform and dielectric features of the surface. For example, in very lossy subsurface media, the amplitude of the signal is attenuated in a shorter time interval, so scanning operations may be limited to the shorter depth. When the objects are deeply buried, some techniques can be used such as “dewow” filtering [1, 2, 9], band-pass filtering [2, 32] and



(a)



(b)

Figure 3.3 Signals obtained from the GPR model: (a) exemplary samples of raw A-scan signals from the test object set at $D = 177$ mm, $P = 175$ mm, and $R = 19$ mm; (b) B-scan image construction for the same test object

time-dependent gain function in other words time-varying gain processing [1, 2, 32].

While the excitation signal is propagating in a dielectric medium, its amplitude decreases depending on the dielectric features of the environment and the propagating path. With the defined parameters, attenuation coefficient [2] is calculated as 0.1088 Np/m. In addition, the environment with the buried object in the center of antennas for lateral position and depths defined at intervals of 40 mm is simulated, and time-varying amplitudes are obtained. For these simulations, a cylindrical PEC object with the radius of 10 mm is used, which is buried at defined geometrical positions. In a particular simulation, the scenario is defined as 238 mm lateral position, 200 mm depth and 10 mm radius. Moreover, as mentioned for GPR model configuration, the antenna system is placed at a distance of 75 mm, and approximate value of 0.66λ [8, 20, 21] (wavelength of guided electromagnetic wave) and very close to ground upper surface. In other words, their distance from the ground is 2 mm [21]. Also, it differs from the conventional radar system due to the short target range and lossy propagation medium for electromagnetic waves [2]. In the proposed model, if there is no buried object in the subsurface, the transmitted field (of the incident field) from the air-ground boundary propagates until the maximum depth of defined surface domain, and when transmitted electromagnetic wave arrives the bottom of the subsurface domain, the amplitude will be zero due to the PML boundary conditions. For this reason, the soil attenuation and the propagation loss, due to the propagation in the dielectric subsurface medium, was demonstrated with the varying amplitudes of the reflected electric fields from the buried object at different depths. In Figure 3.4(a), a Hilbert transformed version of the received signal in the air without subsurface medium and reflected signal from air-boundary ground are represented. This representation is obtained from the Hilbert transform of the signals to utilize the enveloped form of the reflected electric fields. However, background subtracted raw signals are used in the modeling approach and transforming the enveloped form of the signals are used for the explanation of relation between the depth and amplitude change due to soil attenuation.

After the reflection from the upper ground surface, the transmitted part of the excitation signal propagates in the soil environment (subsurface domain), and it attenuates along the propagation path due to the soil dielectric features. In Fig. 3.4(b), sample A-scans corresponding to the mentioned scenarios have been demonstrated to explain the relation between the depth and the magnitude of reflected electric field from the buried object. The principle of decreasing the amplitude of the reflected field as a function of increasing the depth at which the object is buried is used in this work to reduce dimensionality of data through linear regression, more specifically, by extracting the hyperbolic signature. This is the effect of soil attenuation on the

amplitude of the reflected signals. In addition, the amplitude changes of the reflected electric field due to the soil attenuation has been demonstrated in Fig. 3.4(c) by using buried object at different depths, reflection from air-ground boundary and direct wave in the air as well. The variation of amplitudes of the received signals caused by the buried object is explained by using the maximum (peak) amplitude of the enveloped signals. In addition, in Fig. 3.4(d) the ratio between incident and reflected electric fields coming from the buried object at different depths has been shown along with the varying amplitudes of the fields, which indicate the losses.

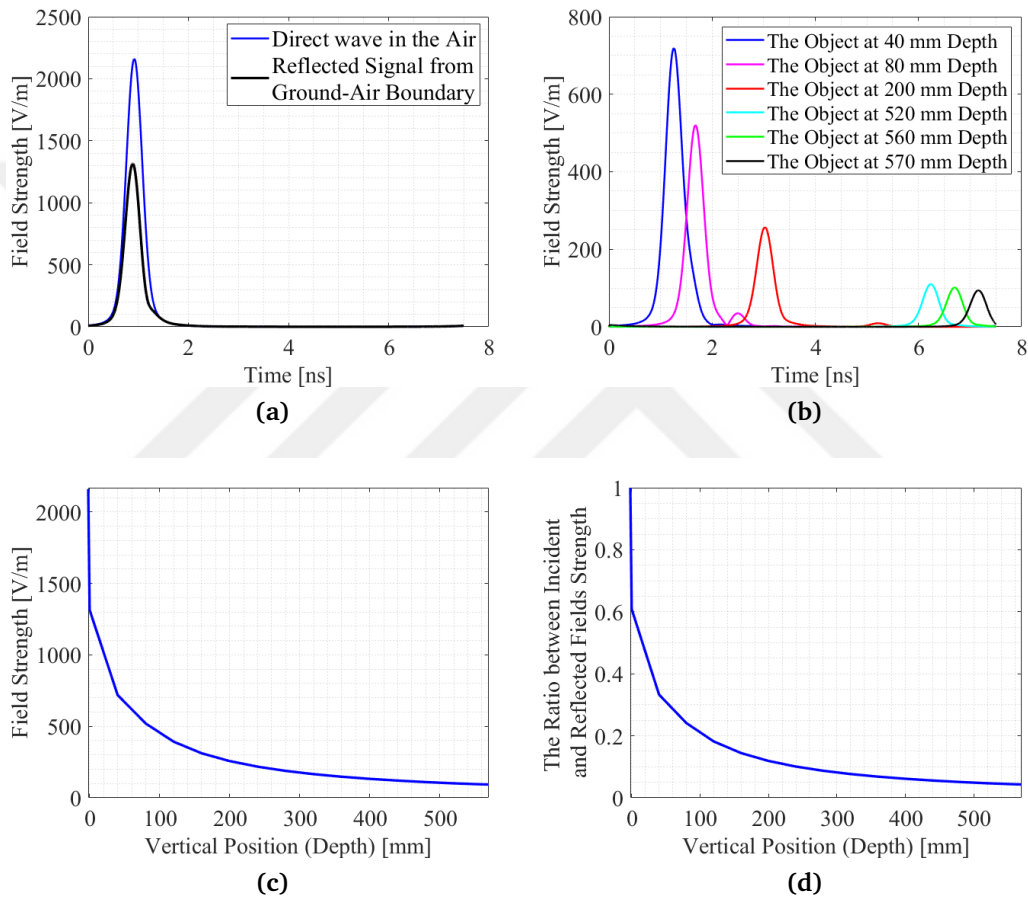


Figure 3.4 Reflected Signals from air-ground boundary and the objects in different depths with their enveloped forms: (a) the received direct wave in the air and reflection from the air-ground boundary; (b) sample scenarios of the A-scan to obtain soil attenuation along the penetration path with the object set at $D = 40$ mm, 80 mm 200 mm, 520 mm, 560 mm, 570 mm, and $P = 238$ mm as well as $R = 10$ mm; (c) amplitude change of the reflected signals versus the depth due to soil attenuation; (d) amplitude ratio between the incident signal and the reflected signals versus the depth

3.3 Arrangement of Data Sets

The A-scans are parameterized using a lateral position of the T/R antenna system; $k = 30$ positions are considered altogether, as indicated in Fig. 3.3. At the same time, each A-scan is sampled at $N_t = 3181$ time steps. This data is aggregated into raw B-scans, in the form of an $N_t k$ matrix that contains the magnitude of the received electric fields strengths in V/m. Thus, the raw B-scan is represented as $E = [E_{ij}]$, with $i = 1, \dots, 3181$, and $j = 1, \dots, 30$ and $E = E_{target} + E_{clutter}$, E_{target} includes reflections from the object and $E_{clutter}$ includes background reflections [5, 59]. The background subtraction operation is applied to reconstruct 2D data (B-scan) as pre-processed B-scan that consists of just reflections from the target [59]. The initial pre-processing is then executed to reduce the data in the time domain by selecting only one out of ten time samples. This leads to compressed B-scans of the size 319×30 .

The problem to be solved is to estimate characteristic parameters of a buried cylindrical object, particularly its depth D , lateral position P , and radius R by using computationally-efficient surrogate modeling approach. According to these characteristic parameters, scenarios are constituted and simulated to obtain B-scans. The pre-processing operation is applied to the collected B-scan data to eliminate the effects of background environment with the effects of other clutters commonly observed in data samples. It is important to elaborate on the structure of the A- and B-scans, as these are critical from the point of view of object characterization. Each A-scan presents two sets of ripples that correspond to the ground and air reflection (the first ripple) and the target reflection (the second ripple). The geophysical parameters of the object but also subsurface permittivity determine both the strength of the reflected field and its time allocation. Further, lateral relocation of the T/R antennas at the along axis leads to creation of a specific pattern of the reflected fields, referred to as hyperbolic signature (hyperbolic feature) [4, 6, 11, 16, 26, 37]. The latter will be exploited in the proposed identification procedure as described in next section.

3.4 Linear Regression-Assisted Surrogate Modeling for Buried Object Characterization

This section outlines a linear regression technique employed to identify hyperbolic signatures of the B-scan data, as well as a novel deep-learning-based approach, Modified Multilayer Perceptron (M2LP) framework used to carry out buried object characterization. The following sections provide the details of both techniques. The programming environment is MATLAB.

3.4.1 Hyperbolic Signature Extraction by Linear Regression Algorithm

When the transmitting and receiving antenna system is moved along the scanning path, a specific pattern is produced as a result reflection from the buried object [4, 6, 11, 16, 26, 37]. It is referred to as a hyperbolic signature, and encodes information about the significant characteristic parameters of the buried object. In this work, the hyperbolic signature is extracted using linear regression [55, 56].

The analytical form of the underlying regression model is a second-order polynomial, which provides a sufficient number of degrees of freedom to represent the object-related information in an adequate manner. The specific data encoded therein is amplitude of the reflected electric field versus the lateral position along the scanning path. A fundamental advantage of this sort of representation is simplification of the data structures being processed by the object characterization framework, both in terms of its sheer amount and dimensionality.

The hyperbolic signature extraction algorithm works as follows. In the first step, the amplitude maxima of each A-scan data acquired along the scanning path are identified for the entire wave travelling time period. The first maximum is the result of reflections from air and the ground upper surface for all A-scan signals, whereas the second maximum emerges from the target reflection. After eliminating the soil effect by applying background subtraction [5–8, 11–17, 20, 21, 59], the estimated hyperbolic signature is created by gathering the pairs of second-highest maxima and the corresponding lateral locations. This data contains information about the buried object position (lateral position P , depth D), but also its radius R . In particular, when radius of the object is increased for at the same depth, the electric field amplitude is also increased. Also, the effects of reflection are observed for a larger number of consecutive A-scan signals around the position where the object is buried, as compared to the object featuring a smaller radius. In other words, the width of hyperbolic pattern increases.

In general form of relation between travelling time and the distance [6, 16, 18, 60] is given in equation (1) indicating two-way travel time. In detail, two-way travel time of the transmitting and receiving signals can be expressed by adding two different times. The reason is that the distance between the antennas and the object center are different from each other. The absolute distance to the object center depends on the distance between the antenna, horizontal and vertical distances between the antennas the object center, but also the object radius. In particular, the antennas are placed on close to the ground surface in the considered model, so that vertical distances of antennas to the object center are neglected.

The specific pattern is constructed from reflections by using the arrival times of the first reflections from the buried object surface for each A-scan in the scanning path. Figure 3.5 shows the geometrical parameters associated with the hyperbolic pattern and its construction from the first reflections due to the buried cylindrical object according to the two-way travel time of the wave are demonstrated. The A-scan index is marked as k and ranges from 1 to 30. Mathematical modeling of the hyperbolic pattern and the times of the first reflections from the object surface is explained with the geometrical configuration and parameters [6, 8, 16, 18, 60] of the data acquisition model. The distance d_t from the object center to the transmitter antenna, and d_r to the receiver antenna, as well as the distance from the transmitter antenna to the starting point at the scanning axis x for each A-scan is defined by calculating the hypotenuse

$$d_{t_k} = ((x_k - P)^2 + D^2)^{1/2} \quad (3.3)$$

$$d_{r_k} = ((x_k - d_a - P)^2 + D^2)^{1/2} \quad (3.4)$$

where d_a is the distance between the antennas (75 mm), P gives the lateral position of the center of object to the starting point, whereas D is the vertical distance between the center of the object and the upper ground surface. The required time t of two-way travelling wave for each A-scan of the first reflection is defined by

$$t_k = \sqrt{\epsilon_r} \frac{(d_{t_k} + d_{r_k}) - 2R}{c} \quad (3.5)$$

where ϵ_r is the relative permittivity of soil medium and c is the light velocity in air. The hyperbolic signature is the merged form of reflected electric fields at time t_k for each A-scan as hyperbolic line and this data is described by

$$HP(k) = [E_1(t_1) E_2(t_2) E_3(t_3) \cdots E_{k-1}(t_{k-1}) E_k(t_k)] \quad (3.6)$$

The assumed analytical form of the function representing the aforementioned pattern is a second-order polynomial of the form

$$y = ax^2 + bx + c \quad (3.7)$$

where x is the lateral position (here, understood as the index of the A-scan), and

y is the electric field strength corresponding to the second-highest maximum of the respective A-scan. The model is extracted based on the pairs $\{x_k, y_k\}$, $k = 1, \dots, N$, with $N = 30$. The least square regression problem is defined as

$$y_k = ax_k^2 + bx_k + c \quad \text{for } i = 1, \dots, N \quad (3.8)$$

which can be written as

$$[a, b, c] \min \left\| \begin{bmatrix} y_1 \\ \vdots \\ y_N \end{bmatrix} - \begin{bmatrix} y_1 = ax_1^2 + bx_1 + c \\ \vdots \\ y_N = ax_N^2 + bx_N + c \end{bmatrix} \right\| \quad (3.9)$$

or, in a matrix form

$$[a, b, c] \min \left\| y - X \begin{bmatrix} a \\ b \\ c \end{bmatrix} \right\| \quad (3.10)$$

where $y = [y_1 \dots y_N]^T$, and

$$X = \begin{bmatrix} x_1^2 & x_1 & 1 \\ \vdots & \ddots & \vdots \\ x_N^2 & x_N & 1 \end{bmatrix} \quad (3.11)$$

The least-square solution to equation (3.11) is given as

$$\begin{bmatrix} a \\ b \\ c \end{bmatrix} = (X^T X)^{-1} X^T y \quad (3.12)$$

Figure 3.6 shows the examples of B-scan data along with their extracted hyperbolic patterns. Herein, a hyperbolic signature extraction is carried out using linear regression with the analytical form of the signature being a second-order polynomial. This methodology is also expressed via using linear regression function as activation function of the hyperbolic line $HP(k)$ which is given by

$$y_k = \phi(HP(k)) \quad (3.13)$$

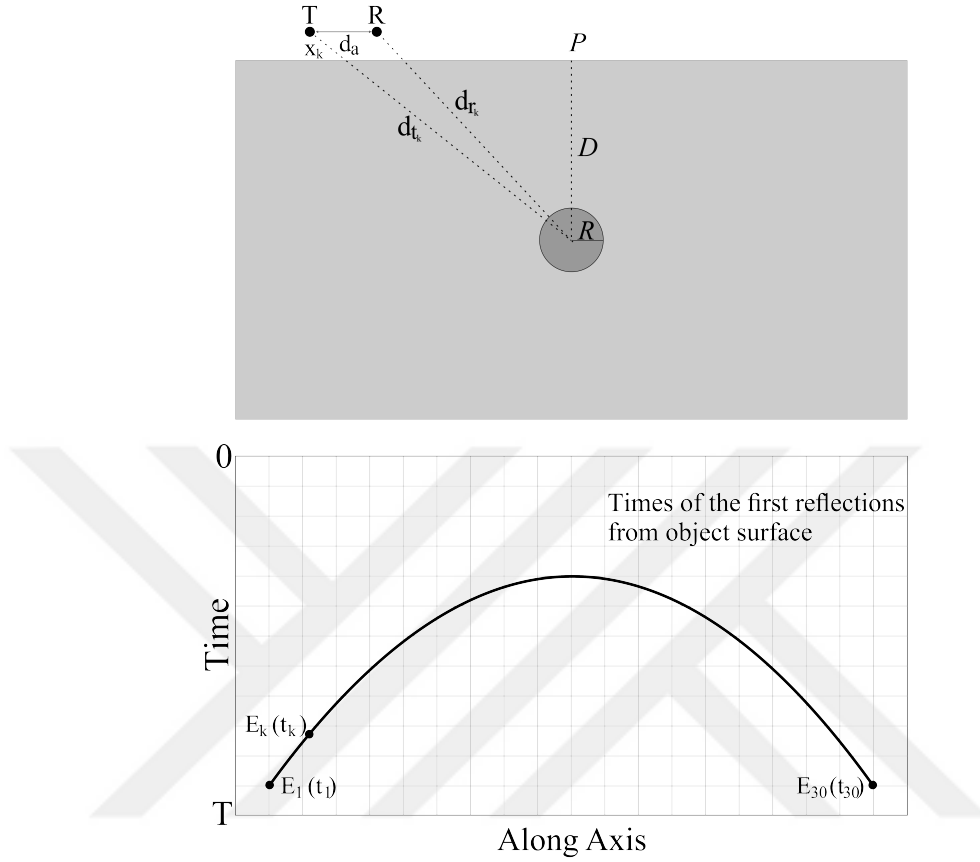


Figure 3.5 The parameters associated with the hyperbolic pattern constructed by the required time of the first reflections from the buried cylindrical object

3.4.2 Deep Learning Based Modified Multilayer Perceptron (M2LP) Framework

Herein, the proposed deep-learning-based modified multilayer perceptron (M2LP) model is introduced for buried object characterization. The flowchart of the modeling process as well as inputs and outputs of the phases of the proposed modeling approach for buried object characterization has been presented in Fig. 3.7.

The proposed framework belongs to a class of MLP networks as it contains fully connected layers. However, it also contains additional layers, normally employed in deep learning frameworks, including a batch normalization layer, Rectified Linear Unit (ReLU) as an activation function. Furthermore, the model is trained using the Adam algorithm [61].

In contrast to the Convolutional Neural Network (CNN)-based regression models, the convolution and pooling layers are not included. In particular, data filtering leads to

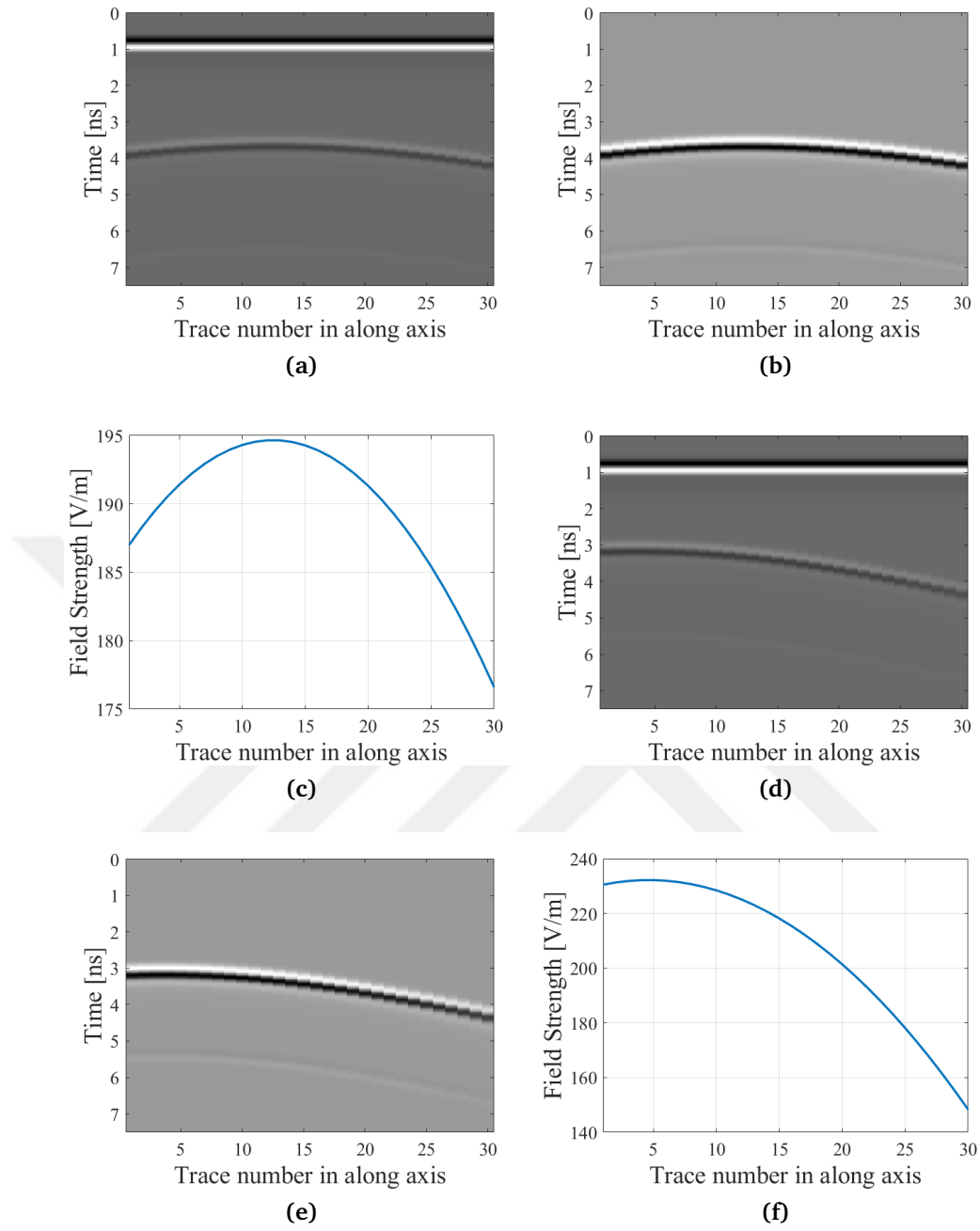


Figure 3.6 Raw and pre-processed B-scan images with their extracted hyperbolic patterns for selected samples concerning the following two test scenarios. The first scenario corresponding to $D = 276$ mm, $P = 172$ mm, and $R = 33$ mm: (a) raw B-scan image, (b) B-scan image with removed background reflections, (c) extracted hyperbolic pattern. The second scenario corresponding to $D = 233$ mm, $P = 83$ mm, and $R = 34$ mm: (d) raw B-scan image, (e) B-scan image with removed background reflections, (f) extracted hyperbolic pattern

unused features remain in the data, also pieces of data losses as the hyperbolic patterns processed in our framework contain the attributes that already carry the knowledge about the characteristic parameters of the buried object. These can be used directly without further pre-processing.

The main architecture of the proposed framework has been shown in Fig. 3.8. It contains of a number of fully connected (FC) layers [58], which operate similarly to feedforward neural networks. Moreover, each FC layer except the last one is followed by the batch normalization (BN) [62] layer, incorporated to eliminate vanishing gradient issues. Further, the activation function employed here is ReLU, which is different from the sigmoid, logsig or tanh used in the traditional neural networks. The last FC layer consists of three neurons, which corresponding to characteristic parameters of the buried object to be estimated. The final layer is a regression layer producing the model outputs.

In the FC layer, neurons realize linear combinations of the input vector entries and adds bias, which is followed by executing the activation function (here, ReLU). The analytical form of ReLU is

$$ReLU(x) = \begin{cases} x & x \geq 0 \\ 0 & otherwise \end{cases} \quad (3.14)$$

As mentioned earlier, the weights within the network are updated using the Adam [61] algorithm, which is a variation of the back propagation method. Batch normalization [62] technique is also utilized to standardize the layer inputs. It also stabilizes the learning process and dramatically reduces the number of epochs required to train the model. Herein, it is used with the batch size of one-tenth of the total number of the training data samples (e.g., 50 for the 500-sample set). The maximum epoch number is 1000, and the data is shuffled in each epoch. Fig. 3.9 illustrates the training progress by providing the loss values versus the iteration number.

3.5 Design of Experiments and Results

For demonstration and verification purposes, the proposed surrogate model has been constructed using the training data sets of different sizes, as explained in Table 3.1. Each sample corresponds to a different scenario concerning the buried object location (lateral position P and depth D), as well as its radius R. The scenarios included in both training and testing datasets are allocated using Latin Hypercube Sampling (LHS) [31].

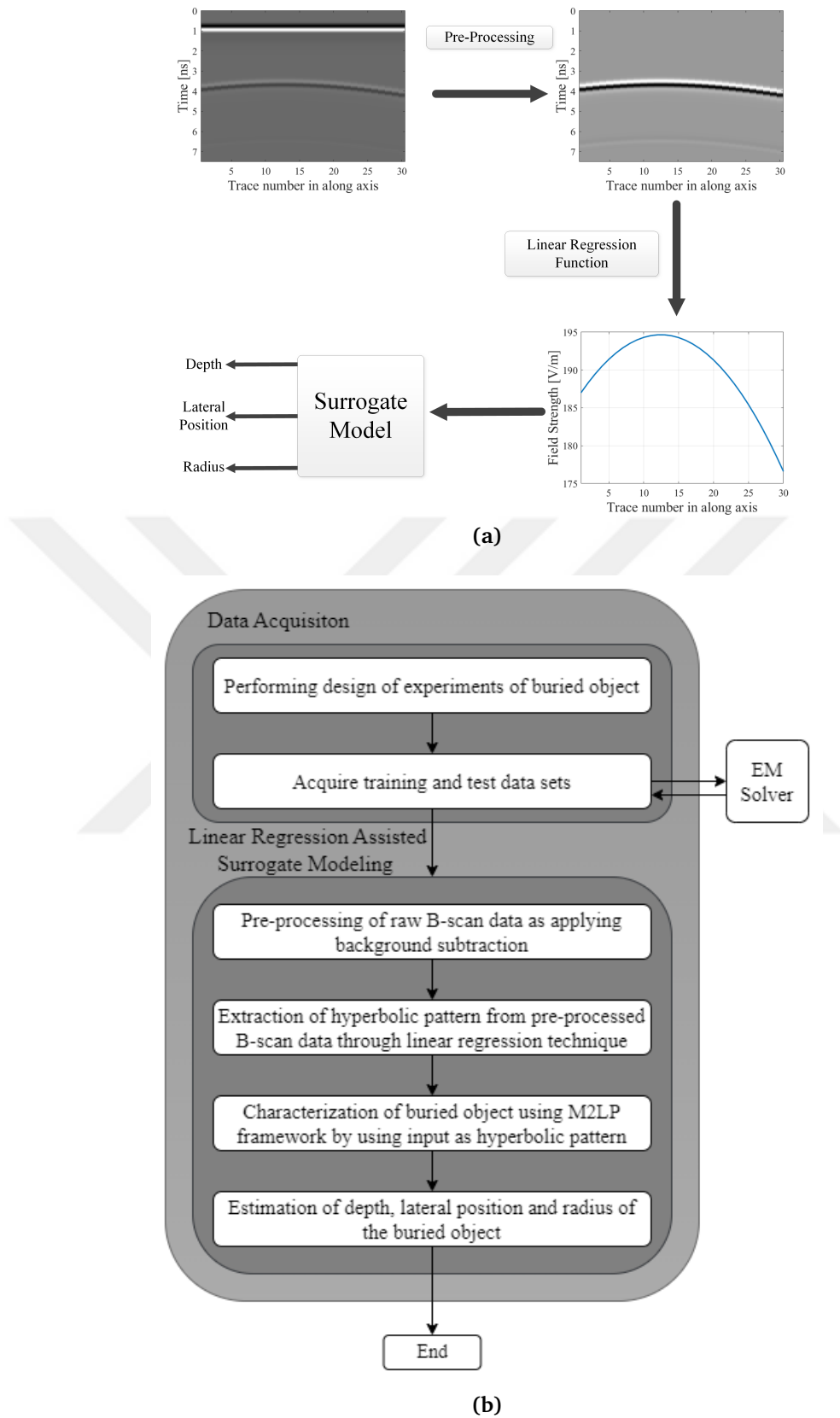


Figure 3.7 Modeling framework: (a) inputs and outputs of the phases in linear regression assisted surrogate modeling, (b) flowchart of the proposed data driven surrogate model for buried object characterization

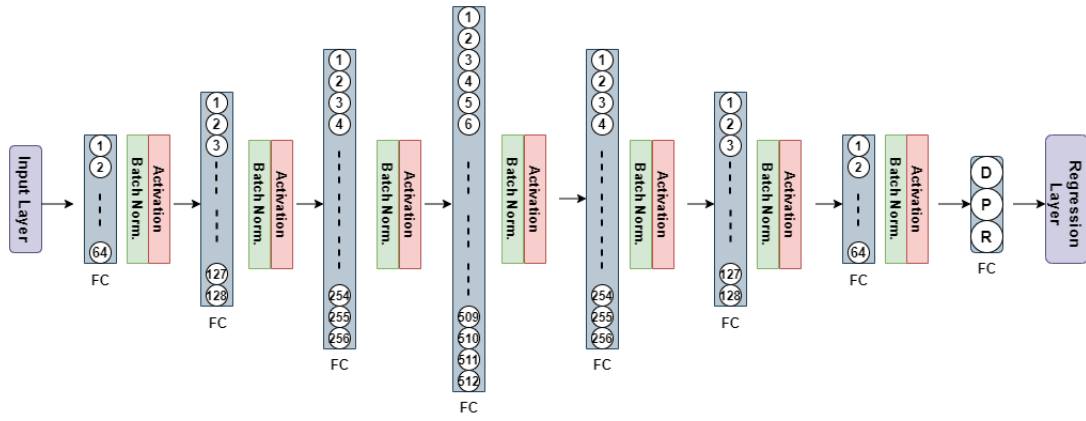


Figure 3.8 Architecture of the proposed M2LP framework

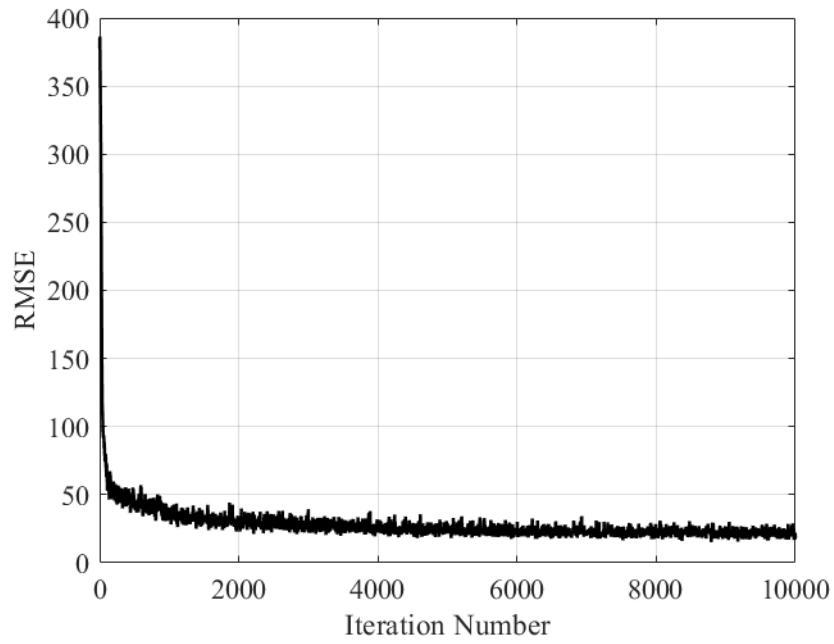


Figure 3.9 The training history of the proposed M2LP model in terms of RMSE (Root Mean Square Error) versus iteration number

Table 3.1 The arrangement of the training/testing data for the considered GPR problem

Data Sets	Total Sample of Scenarios
Training Data Set 1 (DS1)	100
Training Data Set 2 (DS2)	200
Training Data Set 3 (DS3)	300
Training Data Set 4 (DS4)	500
Training Data Set 5 (DS5)	800
Testing Data Set	50
The Dimensions of the B-scans : 319 x 30	

3.5.1 Modeling Results

The predictive power of the surrogate model (both the proposed one and the benchmark) is quantified using the Mean Absolute Error (MAE) and the Relative Mean Error (RME) defined as equations (2.3) and (2.4) respectively.

Table 3.2 shows the MAE and RME for all considered data sets and each characteristic parameter of the object, as well as the average error levels. It can be observed that the accuracy of the proposed framework is satisfactory for practical purposes already with the 500-sample set (DS4), for which the average MAE error falls below ten percent. Consequently, benchmarking of the model will be conducted using this particular dataset.

Fig. 3.10 illustrates some of the geometrical configurations in terms of alignment between surrogate-predicted and target parameters of the object. Table 3.3 provides numerical data, i.e., the target and surrogate-predicted object parameters for selected scenarios. As it can be observed, visual agreement between the actual and predicted object size and location is excellent.

Table 3.2 Prediction performance of considered characteristic parameters for the proposed M2LP framework

Data Set	Characteristic	MAE [mm]	RME [%]	Average MAE [mm]
DS1 (100 Samples)	Depth	43.9	13.3	27.5
	Lateral Position	30.7	22.2	
	Radius	7.8	39.4	
DS2 (200 Samples)	Depth	41.3	11.8	25.1
	Lateral Position	27.4	16.8	
	Radius	6.5	26.2	
DS3 (300 Samples)	Depth	22.6	6.4	13.5
	Lateral Position	14.1	9.8	
	Radius	3.7	15.8	
DS4 (500 Samples)	Depth	14.2	4.4	9.6
	Lateral Position	11.8	7.3	
	Radius	2.6	10.8	
DS5 (800 Samples)	Depth	11.6	3.7	7.6
	Lateral Position	9.0	6.4	
	Radius	2.1	8.7	

As mentioned in the section on configuration of the GPR model, the scanning subsurface dimensions are 400 mm × 600 mm, and the maximum radius of the object is assumed to be 40 mm. The minimum distance between the object and the upper and bottom surface of the soil is defined as 60 mm. With respect to this comment, the effect of reducing the distance to the ground surface is analyzed by adding of a small group

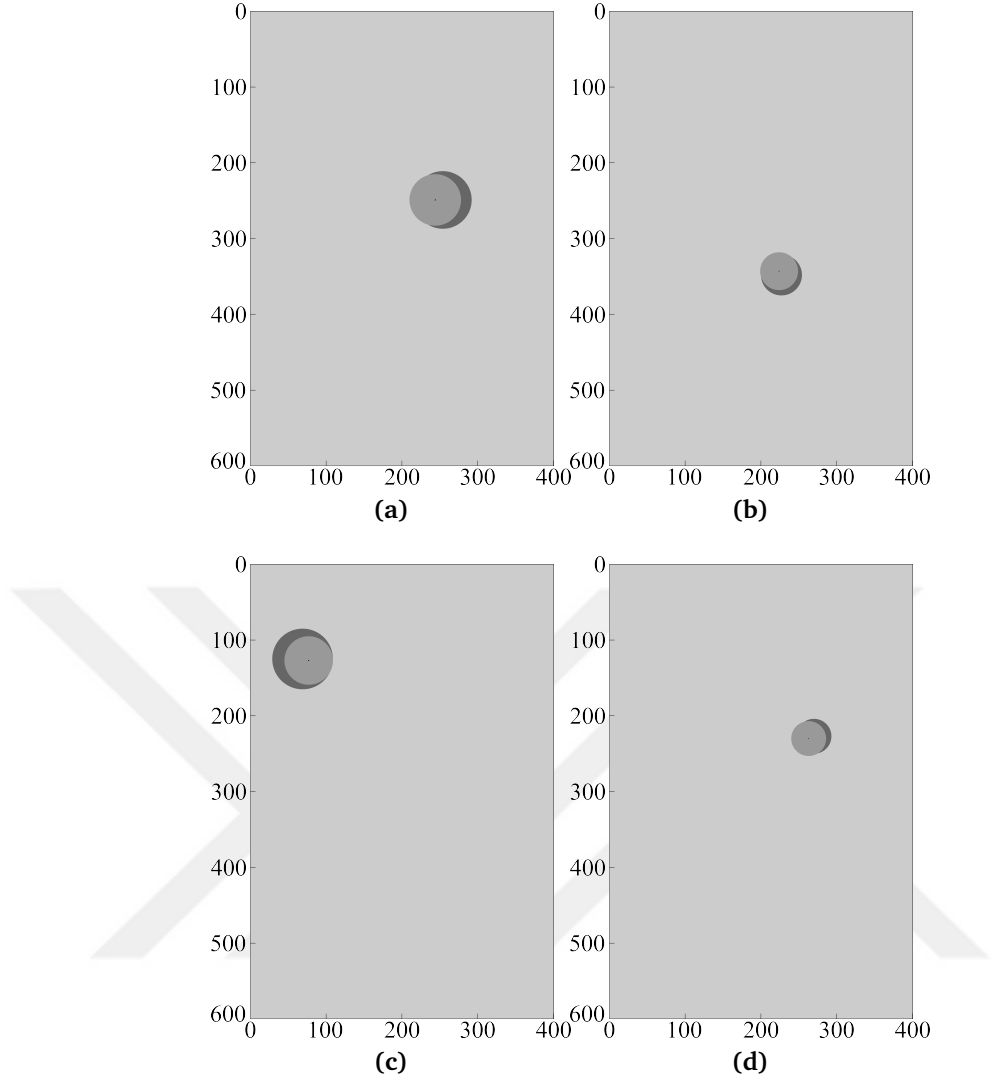


Figure 3.10 Prediction of geophysical parameters of the object, obtained using the proposed surrogate model (built based on the DS4 dataset). The target and surrogate-predicted object marked using the dark- and light-grey shade, respectively: (a) $D = 249$ mm, $P = 254$ mm, $R = 38$ mm, (b) $D = 348$ mm, $P = 227$ mm, $R = 27$ mm, (c) $D = 125$ mm, $P = 69$ mm, $R = 40$ mm, (d) $D = 227$ mm, $P = 270$ mm, $R = 23$ mm

of scenarios (20 samples) that correspond to the object-to-surface distance of 20 mm and 30 mm. The depth characteristic parameter is defined as the distance between the ground surface and the center of the object, so the depth value changes according to the radius value of the object in the scenario. The generated B-scan data for this analysis also can be expressed as $E = [E_{ij}]$, with $i = 1, \dots, 319$, and $j = 1, \dots, 30$ and $E = E_{target} + E_{clutter}$, E_{target} includes reflections from the buried object which is so close to the ground surface and $E_{clutter}$ includes background reflections [5, 59]. The proposed methodology is followed for the new data set to analyze the effects of the object distance to the ground surface. Firstly, the pre-processing is executed by

Table 3.3 M2LP-based estimation of the characteristic parameters in comparison with their true values for selected test scenarios

	<i>D</i> [mm]	<i>P</i> [mm]	<i>R</i> [mm]
True Value	152	102	32
M2LP [this work]	160	101	31
Error	-8	1	1
True Value	187	152	39
M2LP [this work]	187	152	33
Error	0	0	6
True Value	198	146	14
M2LP [this work]	204	143	17
Error	-6	3	-3
True Value	212	318	19
M2LP [this work]	214	311	20
Error	-2	7	-1
True Value	217	201	29
M2LP [this work]	218	198	27
Error	-1	3	2

selecting only one out of ten time samples of the data in the time domain. This leads to compressed B-scans of the size 319×30 . Another phase is applied as the background subtraction operation. The next step is hyperbolic extraction from the pre-processed 2D data. Further, 2D B-scan data is reduced to 1D data consisting of electric field amplitude values along the scanning axis. Finally, the proposed M2LP framework is used for buried object characterization and the results are demonstrated in Table 3.4.

3.5.2 Benchmarking

In this section, the proposed surrogate modelling approach is compared to the latest developed techniques utilized for buried object characterization. The benchmark methods include CNN [12, 13, 26, 37], MLP [3, 7, 9], and SVRM [39]. Also, two of the benchmark cases are analyzed using the M2LP framework operating on different data sets generated by PCA. These methods and cases are briefly explained below.

CNN (Convolutional Neural Network) is a version of deep learning model [6, 12, 13, 16, 26]. The convolutional layer, which is one of the main components of the network, has the ability to automatically extract the data features owing to the convolution filter in the layer. In CNN, several blocks are utilized together, such as a convolution layer (filter), a batch normalization layer, a pooling layer, the activation function, and a fully connected (FC) layer, all involved in the hidden layer. The architecture and the hyper-parameter configuration of the CNN used here is as follows: three convolution layers followed by the batch normalization layer, activation function as

Table 3.4 Characteristic parameter prediction of the proposed M2LP framework assuming short distance between the object and the ground surface

Scenarios and Error Metrics	D	P	R	Closeness to the Ground Surface
True Value [mm]	60	159	40	20
M2LP [this work] [mm]	54	159	35	19
Error [mm]	6	0	5	1
Absolute Relative Error [%]	10.0	0.0	17.5	5.0
True Value [mm]	68	254	38	30
M2LP [this work] [mm]	62	260	34	28
Error [mm]	6	-6	4	2
Absolute Relative Error [%]	8.8	2.4	10.5	6.7
True Value [mm]	43	270	23	20
M2LP [this work] [mm]	53	277	29	24
Error [mm]	-10	-7	-6	-4
Absolute Relative Error [%]	23.3	2.6	26.1	20.0
True Value [mm]	50	344	30	20
M2LP [this work] [mm]	52	328	26	26
Error [mm]	-2	16	4	-6
Absolute Relative Error [%]	4.0	4.7	13.4	30.0

ReLU (Rectified Linear Unit) layer, as well as the two pooling layers included after the last convolution layer, as well as a fully connected layer with three neurons to represent the system outputs. Other user-defined parameters, such as the size and the number of the convolution filter (32, 64, 128) and pooling layer are assigned according to literature recommendations [13]. In addition, it should be mentioned that—in the cited works—the input data is two-dimensional, consequently, the filters of the CNN layers are of the corresponding dimensionality, whereas the filters consisted of the CNN model used for benchmarking in this work are one-dimensional. The CNN model is trained using the Adam algorithm and a batch size of 50. The learning rate has been set to 10^{-3} until maximum epoch number reached to 1000.

Another benchmarking technique is MLP [3, 7, 9, 11, 42]. The model utilized here features the following hyper-parameter configuration: two hidden layers with 32 and 64 hidden neurons, respectively; log-sigmoid activation functions, training by the Levenberg-Marquardt algorithm until the maximum epoch number reaches 1000.

The last benchmark model is Support Vector Regression Machine (SVRM), which belongs to the class of supervised statistical learning techniques [8, 9, 27, 39]. Herein, SVRM has been applied using Bayesian optimization for hyper-parameter adjustment. One of the important components SVRM is a kernel function. Here, it is selected as a Gaussian function to realize nonlinear mapping between the hyperbolic patterns and characteristic parameters of the buried object.

Table 3.5 gathers the performance metrics as well as the training time of the proposed and the benchmark methods. The breakdown of the modelling error for specific characteristic parameters of object (P, D, R) can be found in Table 3.6, whereas Table 3.7 provides prediction performance in concern with comparison of true and predicted geophysical parameters for selected scenarios. It can be observed that the predictive power of the proposed approach significantly better than for the benchmark methods. Both in terms of MAE and RME, the accuracy of our framework is about twice as good as the best benchmark technique (CNN), and about three times better than the accuracy of the remaining methods (MLP and SVRM). Figure 3.11 provides examples of specific scenarios and corresponding object parameter prediction for the proposed approach and the benchmark methods.

Table 3.5 Performance metrics and the training time of the proposed and the benchmark models, averaged with standard deviation over ten independent runs. The models trained using the dataset DS4 (500 samples)

Model	MAE [mm]	RME [%]	Training Time [Minutes]	Processing Time for a Single Input Test Data [ms]
CNN	19.8±1.8	15.4±1.7	6±0.6	10±0.4
MLP	29.6±6.0	26.5±3.4	19±1.0	3±0.2
SVRM	28.8±5.3	24.2±5.0	2±0.3	2±0.2
M2LP [this work]	10.4±1.2	8.1±0.9	4±0.5	7±0.3

Table 3.6 Breakdown of the best prediction performance of the proposed and the benchmark models, over ten independent runs. The models trained using the dataset DS4 (500 samples)

Model	Characteristic	MAE [mm]	RME [%]	Average MAE [mm]
CNN	Depth	25.5	7.9	18.3
	Lateral Position	22.6	16.3	
	Radius	6.6	26.5	
MLP	Depth	19.0	7.3	26.0
	Lateral Position	51.5	35.0	
	Radius	7.3	36.5	
SVRM	Depth	39.7	15.7	23.7
	Lateral Position	23.3	17.5	
	Radius	6.2	30.4	
M2LP [this work]	Depth	14.2	4.4	9.6
	Lateral Position	11.8	7.3	
	Radius	2.6	10.8	

Two of the benchmark cases include different data sets generated using PCA [63–65]. An approach has been proposed for identification of a buried object and to

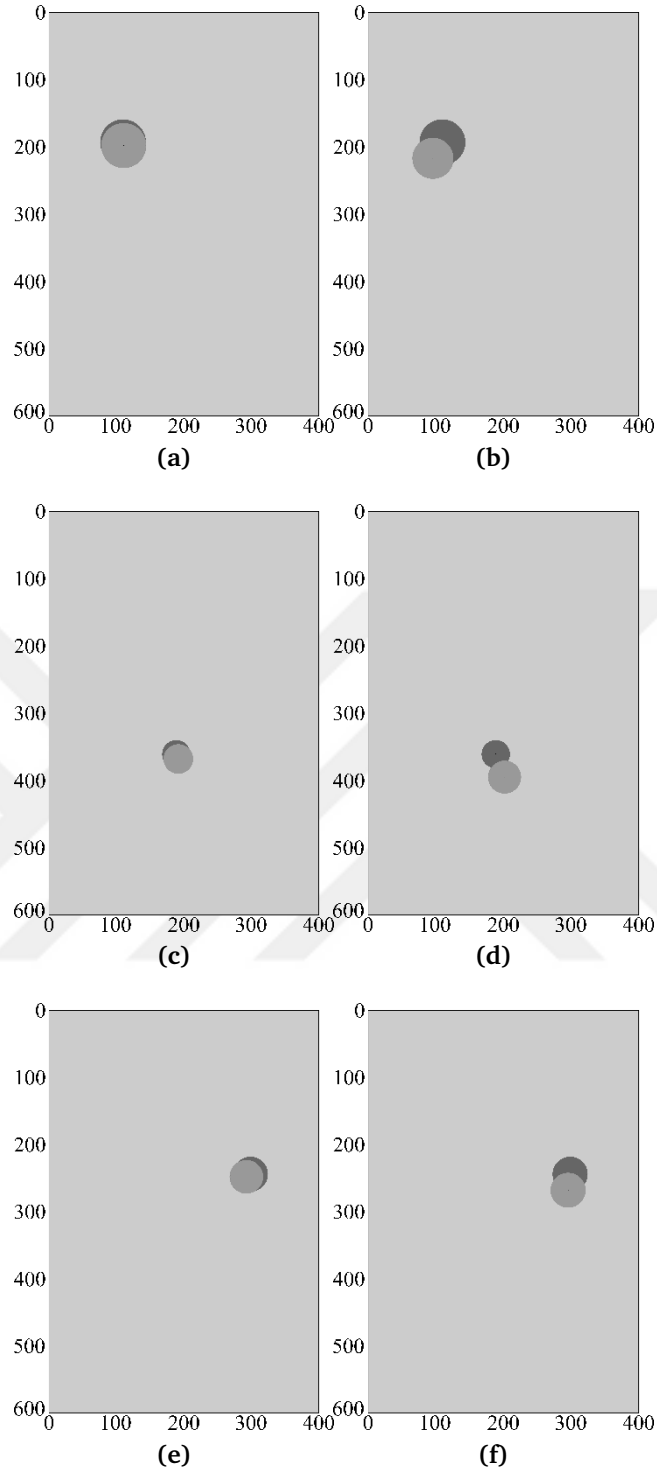


Figure 3.11 Prediction of characteristic parameters of the object, obtained using the proposed surrogate model and the benchmark techniques. The target and surrogate-predicted object marked using the dark- and light-grey shade, respectively:
 Scenario I: $D = 193$ mm, $P = 110$ mm, $R = 34$ mm, (a) proposed model, (b) CNN;
 Scenario II: $D = 361$ mm, $P = 189$ mm, $R = 21$ mm, (c) proposed model, (d) MLP;
 Scenario III: $D = 244$ mm, $P = 299$ mm, $R = 26$ mm, (e) proposed model, (f) SVRM

Table 3.7 Characteristic parameter prediction for the proposed and benchmark methods for selected test scenarios

Model	<i>D</i>	<i>P</i>	<i>R</i>	Error (<i>D, P, R</i>) [mm]
True Value	107	210	33	-
M2LP [this work]	111	208	30	-4, 2, 3
CNN	121	201	24	-14, 9, 9
MLP	113	212	21	-6, -2, 12
SVRM	125	202	32	-18, 8, 1
True Value	279	134	37	-
M2LP [this work]	277	130	36	2, 4, 1
CNN	275	132	30	4, 2, 7
MLP	249	191	26	30, -57, 11
SVRM	248	132	27	31, 2, 10
True Value	170	120	11	-
M2LP [this work]	177	117	14	-7, 3, -3
CNN	186	108	9	-16, 12, 2
MLP	168	135	26	2, -15, -15
SVRM	287	154	25	-117, -34, -14
True Value	287	142	31	-
M2LP [this work]	287	147	31	0, -5, 0
CNN	282	143	26	5, -1, 5
MLP	267	194	26	20, -52, 5
SVRM	282	150	26	5, -8, 5
True Value	260	187	15	-
M2LP [this work]	260	187	16	0, 0, -1
CNN	258	194	13	2, -7, 2
MLP	262	193	26	-2, -6, -11
SVRM	334	169	24	-74, 18, -9

obtain the reflected A-scans by using combined MD (Metal Detector) and GPR sensor [63]. The response features extracted from the peaks and their locations within the subsequent 1D time-varying amplitude signals by means of principle component analysis (PCA), are employed to classify the material type into three groups by using k-nearest neighbor supervised learning classification algorithm [63]. The first case consists of features extracted using PCA similar as in study [63] for the purpose of dimensionality reduction of the B-scan. It is mentioned that B-scan data has the size of 319×30 (319 time steps and 30 A-scans). After applying PCA to raw B-scan, the array of principal components is obtained as the size of 30×29. Its rows are the A-scans and the columns are the components. The principal component array is transformed to 1D feature vector and its size for each B-scan 1×870. In addition, the vertical and horizontal coordinates of the mean and variance magnitudes are extracted from the peak amplitudes and depth indices array. As a result, the input vector is prepared to have the size of 1×874 for the proposed M2LP framework.

The second benchmark case is specialized on clutter reduction by using PCA [64, 65], where only the reflections coming from the buried object are obtained and linear regression technique is employed to identify hyperbolic signatures of the B-scan data. After that, the proposed deep-learning-based approach, M2LP framework is used to carry out buried object characterization. In Fig. 3.12, for a sample test scenario the principal components as an image (in 2D data form), and a B-scan image pre-processed (clutter reduced) using PCA are demonstrated. In addition, the last benchmark case including a study of characterization of geophysical parameters with A-scan analysis [66]. The performance results of that study which is computationally efficient surrogate modeling via a novel deep learning-based framework that focuses on the object characterization in terms of its geophysical parameters with A-scan analysis [66] and by using raw data (without any background subtraction operations) are added. These benchmark cases only have approximate training and testing time durations, so this data is not considered in the comparative study. In Table 3.8, performance metrics of the proposed and the benchmarking cases, in the form of the average values and standard deviations of MAE and RME over ten independent runs have been represented.

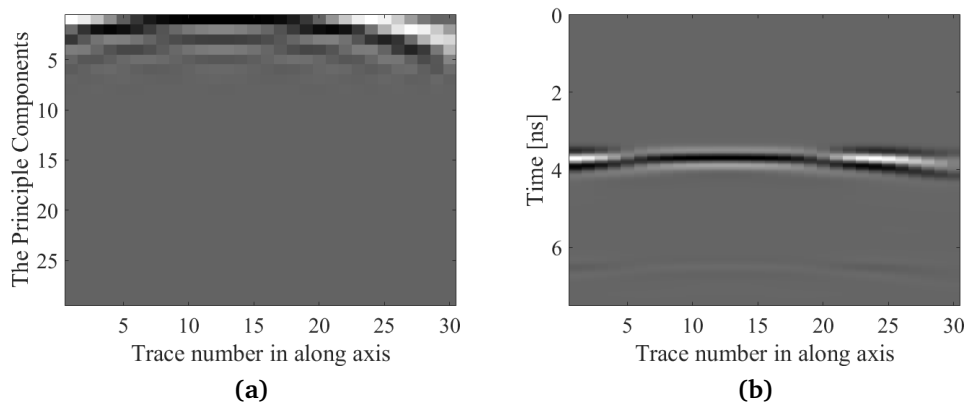


Figure 3.12 A sample test scenario corresponding to $D = 276$ mm, $P = 172$ mm, and $R = 33$ mm: (a) extracted principle component features by using PCA; (b) pre-processed (clutter reduced) B-scan image by using PCA

The results indicate that linear regression assisted hyperbolic signature approach with the proposed deep-learning-based M2LP framework features smaller error as compared to other cases including different data sets with the proposed framework and different method (A-scan analysis) with TFRM framework [66]. It can be observed that, in a qualitative sense, according to the presented results the proposed methodology is superior to the benchmark cases.

Table 3.8 Average performance metrics of the proposed and the benchmark cases and standard deviation, computed over ten independent runs

Model & Methodology	MAE [mm]	RME [%]
M2LP [this work]	10.4±1.2	8.1±0.9
M2LP with hyperbolic signature, reconstruction B-scan via PCA	17.4±0.4	20.4±0.4
M2LP with extracted features, the principal components and mean, variance coordinates	17.2±0.9	21.3±0.9
TFRM with unprocessed raw 1D data, A-scan analysis [66]	12.1±0.1	16.5±0.5

3.6 Data-Driven Surrogate Modeling with Noisy Data for Validation of Realistic Scenarios

This section addresses object characterization assuming more realistic scenarios, namely, noisy data. For demonstration purposes, new data sets were created by adding random noise to the A-scan samples from the DS4 dataset. This is to emulate the environmental and internal noise of the data gathering system in GPR the model, and to evaluate its effects on the proposed modeling methodology. The selected scenarios from the testing data have been shown in Fig. 3.13.

The literature offers different approaches to noise incorporation [16, 24, 25, 37]. The background noise of real data has been integrated with the buried object reflections (B-scans) of synthetic data [16, 24]. In a study [37], a noisy data set has been generated by adding randomly replaced black and white pixels on B-scans. This allows for investigating different noise levels in synthetic data generated using the gprMax simulation tool to predict diameter of the target on the extracted hyperbola from the pre-processed B-scans [37]. In this work, the supplementary noisy data sets were created with different signal-to-noise (SNR) levels of 20 dB and 30 dB, by adding the white Gaussian noise [25] to the data set including 500 samples (dataset DS4). This enables emulating conditions that are closer to the realistic situation or on-site applications.

For this investigation, the proposed M2LP framework has been compared with all benchmark models (CNN, MLP, SVRM). The results are presented in Table 3.9 and Table 3.10 provides a comparison of model-predicted versus actual characteristic parameters for selected test cases. The results indicate that the proposed M2LP framework outperforms the benchmark by a considerable margin. In particular, the average MAE of M2LP is lower by a multiplicative factor of 1.6, 1.9 and 1.9 as compared to CNN, MLP, and SVRM, respectively for SNR = 30 dB, whereas the

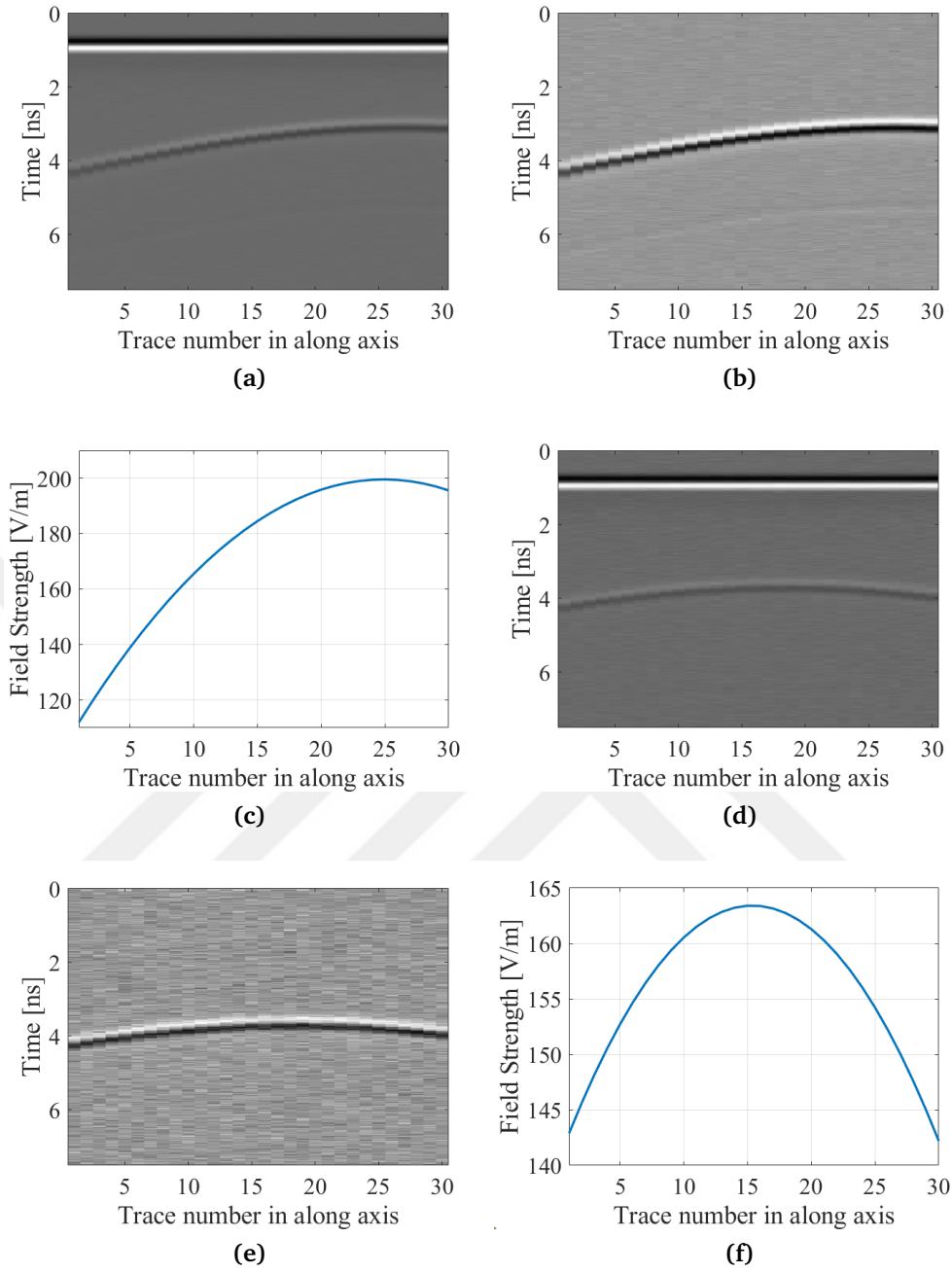


Figure 3.13 Raw and pre-processed B-scan data along with their extracted hyperbolic patterns for selected samples from the test scenarios including noise. The SNR for first scenario is 30 dB, whereas the objective parameters are $D = 212$ mm, $P = 318$ mm, and $R = 19$ mm: (a) raw B-scan image, (b) B-scan image with removed background reflections, (c) extracted hyperbolic pattern. The SNR for the second scenario is 20 dB, $D = 259$ mm, $P = 231$ mm, and $R = 12$ mm: (d) raw B-scan image, (e) B-scan image with removed background reflections, (f) extracted hyperbolic pattern

improvement is as high as 1.5, 1.8, and 1.6 over CNN, MLP and SVRM, respectively, for SNR = 20 dB.

Table 3.9 Prediction accuracy of characteristic parameters of the buried object for all models. Considered scenarios include noisy data with SNR of 30 dB and 20 dB

Model	Characteristic	SNR = 30 dB			SNR = 20 dB		
		MAE [mm]	RME [%]	Average MAE [mm]	MAE [mm]	RME [%]	Average MAE [mm]
CNN	Depth	38.1	14.1	28.9	58.1	19.7	37.9
	Lateral Position	43.0	27.6		48.9	30.8	
	Radius	5.5	25.2		6.7	31.4	
MLP	Depth	53.7	18.2	35.6	72.9	27.5	45.6
	Lateral Position	45.8	30.0		56.6	37.2	
	Radius	7.1	36.3		7.4	36.6	
SVRM	Depth	47.8	18.9	35.8	61.5	24.1	41.1
	Lateral Position	54.3	37.0		56.5	37.1	
	Radius	4.6	21.6		5.4	26.8	
M2LP [This work]	Depth	25.3	7.8	18.8	28.4	9.6	24.5
	Lateral Position	26.6	17.6		39.3	24.5	
	Radius	4.6	21.2		5.8	27.7	

3.7 Data-Driven Surrogate Models for Buried Object Characterization via Measurement Data

For the sake of supplementary demonstration of the proposed surrogate modeling approach, an additional study has been conducted using experimental data (B-scans) collected through the measurements in a “sand pool” environment. A sparse data set is utilized that contains 33 scenarios in total, 27 scenarios for training and the remaining scenarios utilized for testing. It should be emphasized gathering the experimental data is an expensive endeavor due to the considerable manual labor involved (digging and burying targets with high accuracy), as well as the adjustments/maintenance of the measurement system. These are the main reasons for using data-driven surrogate modeling approaches in the field of GPR characterization/detection systems. In this section, our aim is to demonstrate that the proposed technique is also applicable in the case of using physical measurements as a source of data. The experimental samples are obtained in the laboratory at Yıldız Technical University. During the process, raw B-scan data corresponding to various scenarios are generated by the impulse ground

Table 3.10 Predicted characteristic parameters of the buried object versus true values for selected test scenarios

Model	<i>D</i>	<i>P</i>	<i>R</i>	Error (<i>D, P, R</i>) [mm]
True Value	170	120	11	-
M2LP [30 dB]	182	124	11	-12, -4, 0
CNN [30 dB]	196	114	17	-26, 6, -6
MLP [30 dB]	195	102	27	-25, 18, -16
SVRM [30 dB]	289	158	18	-119, -38, -7
True Value	372	196	35	-
M2LP [30 dB]	361	195	33	11, 1, 2
CNN [30 dB]	302	193	29	70, 3, 6
True Value	372	196	35	-
MLP [30 dB]	380	160	25	-8, 36, 10
SVRM [30 dB]	323	199	28	49, -3, 7
True Value	276	172	33	-
M2LP [20 dB]	265	161	28	11, 11, 5
CNN [20 dB]	209	180	20	67, -8, 13
MLP [20 dB]	307	200	24	-31, -28, 9
SVRM [20 dB]	288	189	28	-12, -17, 5
True Value	143	333	25	-
M2LP [20 dB]	150	325	22	-7, 8, 3
CNN [20 dB]	201	216	23	-58, 117, 2
MLP [20 dB]	130	316	28	13, 17, -3
SVRM [20 dB]	247	274	27	-104, 59, -2

penetrating near-zone radar system, which is utilized in various subsurface imaging operations [50–53]. Figure 3.14 shows the experimental setup. The measurements are taken in a wooden pool filled with dry soil utilized as subsurface. The scanning subsurface domain has the dimensions of 1.90 m (width), 0.22 m (depth) and 1.15 m (length). The cylindrical PEC object is buried in the inhomogeneous dry soil consisting of a mixture of small stones and sand. The experimental setup (GPR, transmitter and receiver antennas) is manually relocated above the subsurface along the scanning path of the approximate length of 1.90 m. Each B-scan is set as 255 (discrete time step) $\times$ 95 (A-scan number). The discrete time step is arranged for 6 ns according to the sample rate within 7.5 ns of the compressed data obtained by using gprMax simulation tool. In addition, the measurement data in the form of $N_t \times k$ (position number at the scanning aperture) matrix includes 255 $\times$ 95 elements. The value of the A-scan number is expressed as 95 and it is defined according to the step size which is 20 mm gives the distance between two consecutive A-scan measurements [8, 10, 20].

According to the proposed approach, background subtraction is applied to obtained raw B-scan images. Subsequently, hyperbolic signatures are extracted from pre-processed measured data (B-scans) by using the linear regression technique.



Figure 3.14 Illustration of the GPR configuration utilized for generating experimental training and testing data used by the proposed surrogate-assisted buried object characterization framework

Figure 3.15 shows an example scenario with its raw B-scan data, pre-processed B-scan data, and the extracted hyperbolic signature utilized as the input for the purpose of surrogate modeling. By following this approach, 1-D reduced data set is obtained. The extracted hyperbolic pattern, representing the amplitude of the reflected impulse versus the A-scan index along the scanning path is associated with the characteristic parameters of buried object in terms of depth, lateral position and radius.

Table 3.11 shows the modeling performance of the proposed surrogate approach and one of the benchmark methods, CNN. The error of the M2LP framework is 27.3 mm (MAE) and 17.9 % (RME), which is the average of considered outputs (depth, lateral position and radius). The CNN model exhibits considerably worse accuracy with the average error being almost twice as high. Figure 3.16 shows the prediction performance for selected scenarios. It should be noted that the proposed method accurately predicts the location of the object, whereas CNN prediction is poor with this respect. The prediction of the object size is considerably better for the M2LP model as well.

Table 3.11 Prediction performance of considered characteristic parameters for M2LP and CNN models by using real data set obtained from measurements

Model	Characteristic	MAE [mm]	RME [%]	Average MAE [mm]
M2LP	Depth	29.3	27.5	27.3
	Lateral Position	47.9	5.3	
	Radius	4.5	21.1	
CNN	Depth	18.0	15.2	43.0
	Lateral Position	105.9	11.3	
	Radius	5.2	23.0	

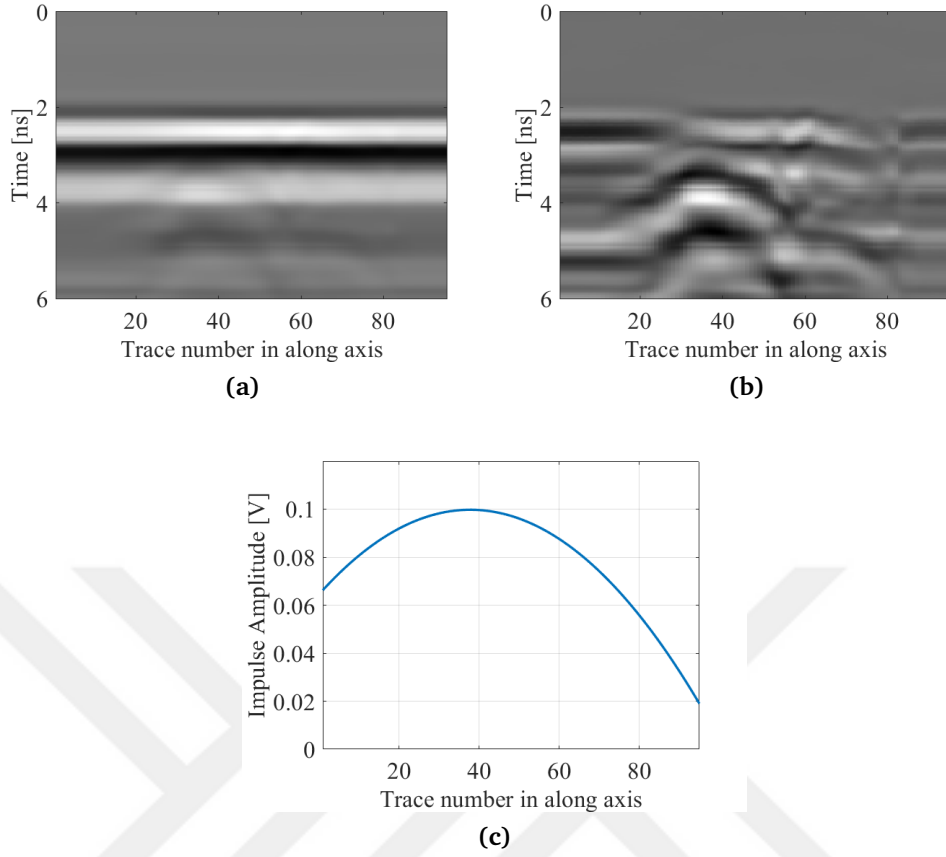


Figure 3.15 Raw and pre-processed B-scan data with their extracted hyperbolic patterns for selected sample concerning the following test scenario which corresponds to $D = 750$ mm, $P = 90$ mm, and $R = 15$ mm: (a) raw B-scan image, (b) B-scan image with removed background reflections, (c) extracted hyperbolic pattern

3.8 Conclusion of Buried Object Characterization via Data-Driven Surrogates and Regression-Enabled Hyperbolic Signature Extraction

In this chapter, a novel regression-assisted surrogate modeling technique is introduced for characterization of buried object. The proposed M2LP framework consists of deep-learning-based feedforward network configuration in the form of a series of blocks of fully connected layers, batch normalization, and activation layers, the latter involving the ReLU activation functions. The B-scan data utilized for object characterization is obtained by using FDTD-based electromagnetic simulation toolbox of a Ground Penetrating Radar (GPR) for estimating of geophysical parameters of a cylindrical shape objects of various radii, buried at different locations. One of the novelties of the work is utilization of the hyperbolic signatures, extracted from the pre-processed B-scans by using linear regression technique, as the inputs of the modeling process. The latter simplified and reduces dimensionality of the

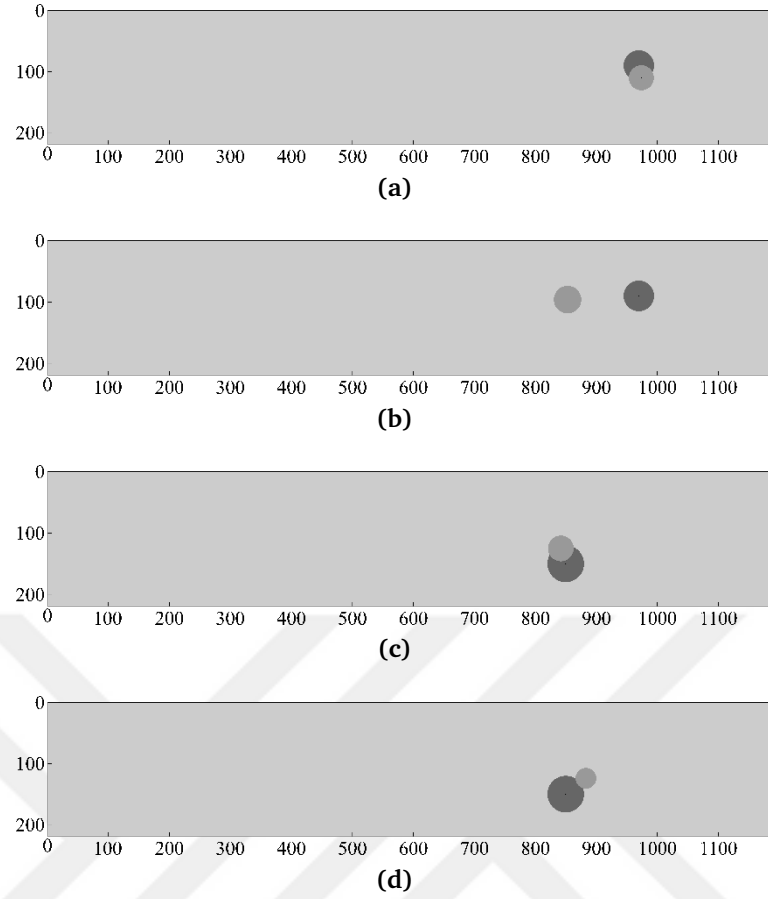


Figure 3.16 Prediction of geophysical parameters of the object, obtained using the proposed surrogate model and the benchmark model, CNN (using measured data). The target and surrogate-predicted object marked using the dark- and light-grey shade, respectively: Scenario I: $D = 90$ mm, $P = 970$ mm, $R = 25$ mm, (a) proposed model, (b) CNN; Scenario II: $D = 150$ mm, $P = 850$ mm, $R = 30$ mm, (c) proposed model, (d) CNN

input parameter space. This does not only enhance computational efficiency of model construction but also permits a rendition of accurate data-driven surrogates using small number of training data samples. The presented approach is compared to a number of benchmark models including MLP, SVRM, and CNN. The results indicate competitive prediction performance of our method with the average MAE as low as 10 mm versus 18.3 mm for the best benchmark model, CNN. For supplementary verification, the proposed methodology has been applied to buried object characterization using noisy data sets of SNR levels corresponding to 30 dB and 20 dB. The modeling error in these cases corresponds the average MAE of 18.8 mm and 24.5 mm respectively, which is excellent given the scarcity of the datasets. Finally, modeling based on experimental data has been discussed as well, corroborating the relevance of the presented approach. For the considered scenarios, the prediction accuracy of the buried object parameters has been shown to be almost twice as good as the best benchmark method (again, CNN). The numerical and experimental data

gathered in the paper is indicative of the competitive accuracy of the introduced methodology, as well as the relevance of the considered algorithmic components, especially dimensionality and problem complexity reduction through delegating the modeling process to hyperbolic signature level. The approach presented in this chapter can be considered a viable alternative to existing surrogate-assisted techniques for buried object characterization both with respect to accuracy and computational efficiency.



COMPREHENSIVE ANALYSIS OF BURIED OBJECT CHARACTERIZATION WITH SEVERAL VARIED DATA SETS AND DATA-DRIVEN SURROGATE MODELS

4.1 Introduction to Characterization Analysis of Buried Object via Several Varied Data Sets and Data-Driven Surrogate Models

The mentioned techniques in literature review apply pre-processing, clutter reduction methods for data identification and characterization of a buried object. It is emphasized that one of the most common method is background subtraction and removing air-ground surface echo. Some other methods are principle component analysis (PCA) [5, 64, 65, 67], morphological component analysis (MCA) [68], singular value decomposition (SVD) [5, 67], independent component analysis (ICA) [5, 67], also ICA with multifractal spectrum [35] for clutter reduction in other words elimination of direct wave, reflections from air-ground surface and subsurface medium, also noise suppression. However, these applications increase computational cost and the techniques of feature extraction are used for reduction in dimensionality of data to obtain computational efficiency.

In addition to clutter removing, another pre-processing application of PCA is feature extraction for the aim of dimensionality reduction of B-scan image (2D data) [27, 63]. Also, feature extraction techniques are applied on the average window subtracted B-scan images [27] for material classification object using support vector machine (SVM) and NN classifiers. The buried objects are identified using geometrical features such as minor and major axes, along with statistical features (SFs) obtained via principal components (PCs), mean, variance and kurtosis, etc. [27]. In similar, the identification of buried materials using SVM and extracted statistical features (SFs) from fractional domain-envelope curve [69]. In another study, performances of the extracted statistical features from the reflected signals in noisy environment like variance, deviation, kurtosis and skewness for classification material type are

compared with the spectral features [70]. Furthermore, the extracted features from reflected wave in time and frequency domain have been used for discrimination of air, water, conductor materials in classification approach and estimation of depth and radius in regression approach by using k-Nearest Neighbor (k-NN) model [71].

In this chapter, the common used data type of reflected received signals as B-scan (2D data), data including extracted features from raw time signals as PCs and SFs (mean, variance, standard deviation, skewness and kurtosis) and time frequency spectrogram (TFS) of consecutive A-scans (1D data) are used for buried object characterization. Also, data diversification approaches are followed by producing new data sets from raw time signals to propose as an alternative to image processing, hyperbola investigations on B-scan images and to analyze estimation performances of extracted SFs. It can be evaluated as a contribution that the characterization of buried object is achieved by using data-driven surrogate model constructed with limited linear sampling scenarios, without removing ground reflections, background subtraction operations, B-scan image processing, and hyperbola investigations (unprocessed). In particular, removing ground reflections or background subtraction operations might be challenging for buried object scenarios that belong to more than one soil type or host medium dielectric features, which changes with the percentage of the water content. Herein, four different subsurface media according to their percentage of the water content [28, 29] are taken into account to increase and analyze the effect of complexity arising from the environments with different dielectric permittivity.

The main aim of this analysis is to prioritize computationally efficient surrogate modeling studies and propose a novel deep regression network (DRN) that focuses on the estimation of characteristic parameters independently from each other and simultaneously with TFS data without any application of clutter reduction method by using a few linear sampled train scenarios. In addition to proposing a novel framework (DRN) that employs the aforementioned a unique methodological approach, the performance comparison with the state of the art techniques reported in the literature is provided for parametric estimation of buried object characterization which is applied independently based on the subsurface dielectric properties.

4.2 Configuration of GPR Model and Scenarios of Data Acquisition

The main problem to be handled in this work is estimation of the characteristic parameters of buried cylindrical object simultaneously and independently from each other. In particular, the characteristic parameters to be estimated are defined as depth D , lateral position P and radius R . The geometrical explanation of the characteristic

parameters has been provided in Figure 4.1.

In study cases, data sets are prepared by using unprocessed, raw B-scan data obtained from the GPR model to be discussed in the next sub-section. Parametric estimation process of the characteristic parameters in the formulated problem is applied via deep regression network framework for time-frequency spectrogram (TFS) of consecutive A-scan data (1D time-varying power amplitude signals).

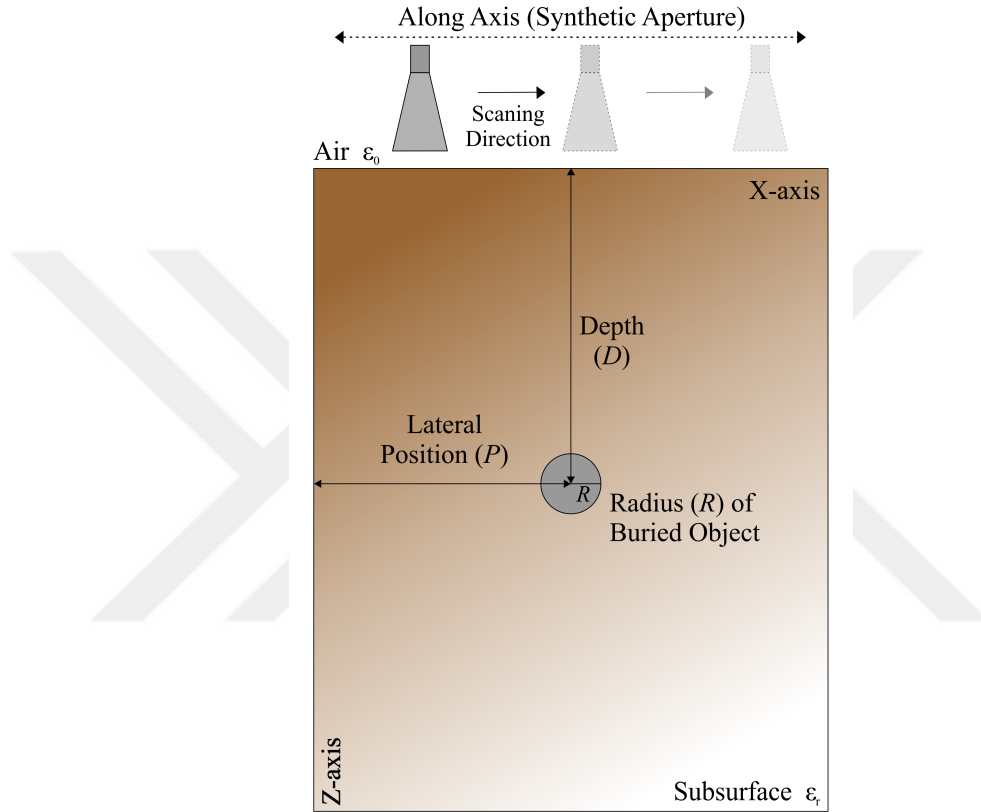


Figure 4.1 The two-dimensional model diagram explaining the geophysical parameters to be solved

Herein, a ground penetrating radar (GPR) problem is solved by characterizing a buried cylindrical PEC object in terms of localization in detail vertical, lateral (horizontal) position and object size as radius. This characterization problem is analyzed by using data-driven surrogate models. For data acquisition, buried object characterization model is evaluated via 3D full-wave electromagnetic (EM) analysis tool and a C-band pyramidal horn antenna is used in a monostatic configuration [10]. The used GPR antenna is placed very close to air-ground surface. The transmitted signal is defined as a modulated Gaussian signal with a center frequency of 6 GHz. The power amplitude of the received reflected signals varies during traveling time of wave penetrated at the defined subsurface depth which is defined as 12 ns by using equation (2.1). 1D signal, A-scan is a time-varying normalized power amplitude signal obtained at the one point

along the synthetic aperture. In other words, one A-scan length is 600 (time steps). B-scan image constructed by concatenating the A-scans (16 A-scans) obtained as a result of reflections of the transmitting excitation signal along a selected scanning path.

The scatterer is modelled as cylindrical perfectly electric conductor (PEC) object such as a wire, a pipe, or a rebar. The subsurface domain dimensions are defined as $400 \times 300 \times 500$ mm. The subsurface is assumed to be a dispersive subsurface medium and its dielectric features is described by the extended Debye model [30] according to the percentage of water content described with equation (2.2). In addition to data set of characterization studies in Chapter 2 one more soil type is utilized. For this model, four different types of soil media are chosen according to their water content of 0.2 %, 2.8 %, 5.5 % and 6.2 % [30]. Their parameters are as ϵ_s and ϵ_∞ relative permittivity values at zero and infinite frequency, respectively, t_0 the relaxation time, σ the conductivity are demonstrated in Table 2.1 and Table 4.1. The parameter space of scenarios has four dimensions, which include the depth, lateral position, radius, and the soil water content. The train data set is designed with a linear sampling method which is used for generating 420 sample points within the given parameter ranges and it is presented in Table 4.2.

Table 4.1 Subsurface dielectric features of Extended Debye model according to water content [30]

WC %	ϵ_s	ϵ_∞	t_0 [ns]	$\sigma [\Omega^{-1}m^{-1}]$
6.2	9.14	5.93	0.8	6.7×10^{-3}

For the test data set, Latin Hypercube Sampling (LHS) [31] method which is statistical method generating a near-random sample of parameters from multidimensional distribution is used for generating randomly selected sample points of the GPR model consisting of four type soil media. Total of 81 scenarios are generated to be used for performance evaluation of surrogate models. In Figure 4.2, the configuration of the characterization parameters for the training and test data sets are demonstrated.

Table 4.2 The arrangement of the training data for the considered GPR problem

Parameter	Sample range	Step size
Depth of Object [mm]	100-400	50
Lateral Position [mm]	140-280	70
Radius of Object [mm]	10-50	10
Water Content[%]	0.2, 2.8, 5.5 and 6.2	
Number of A-scans in each scenario	16	
Total Number of Scenarios	420	

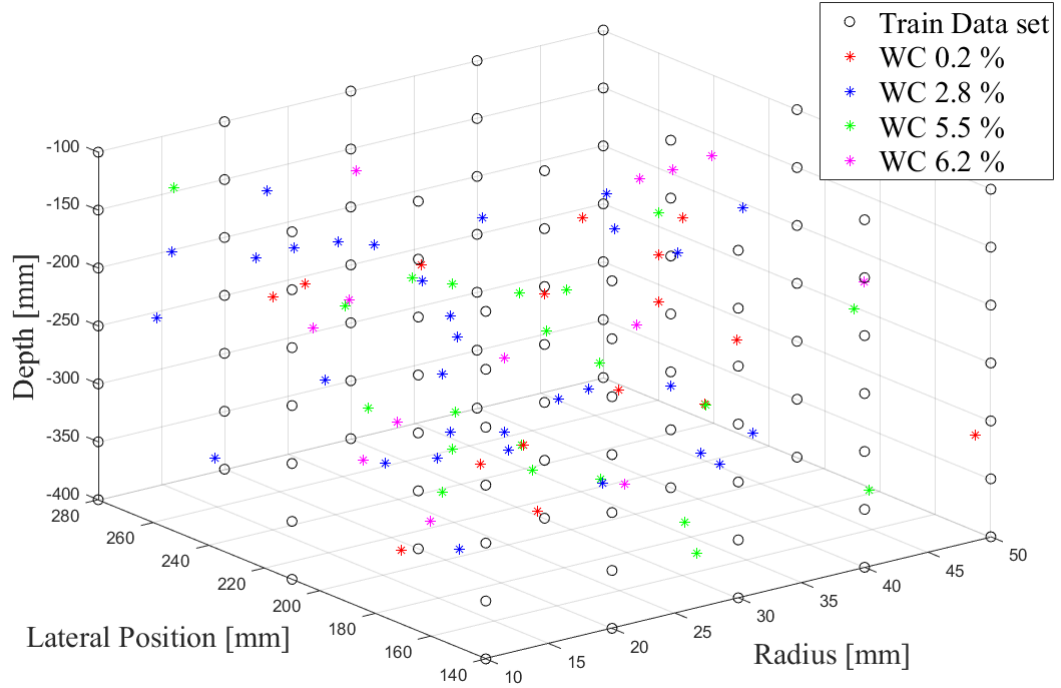


Figure 4.2 Configuration of characterization parameters for scenarios in the training and testing data sets in 3D plot demonstration with regards to depth, lateral position and radius

4.3 Study Cases for Characterization Analysis

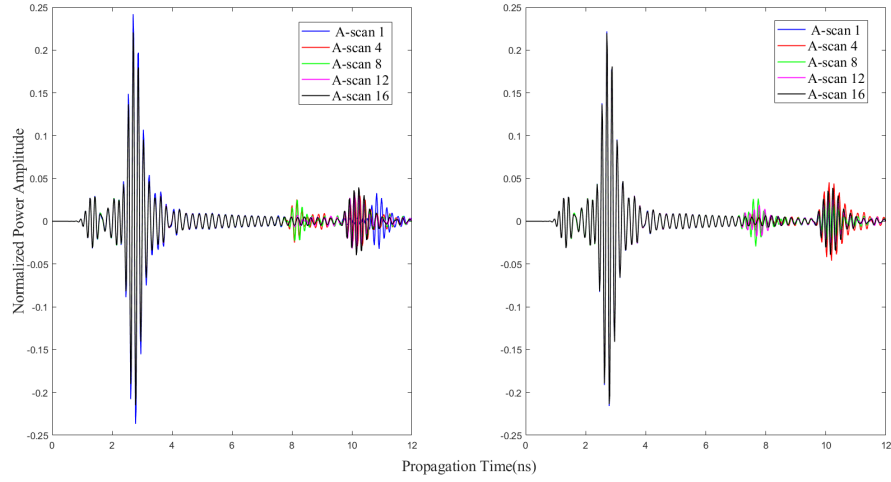
In the studied GPR problem, received reflected power amplitude signal with three different study cases are analyzed for using inputs of AI based surrogate models to estimate the aimed characteristics of the buried object. In literature, to solve inverse GPR problem B-scan image (2D data), extracted features and A-scan signal are used. Herein, these data types are analyzed with three case study which are taken into consideration for data sets as Case 1– 2D of raw time signal, Case 2– extracted features from raw time signals and Case 3– 1D raw A-scan signal.

4.3.1 Case 1- Data Set of Raw Time Signals in 2D

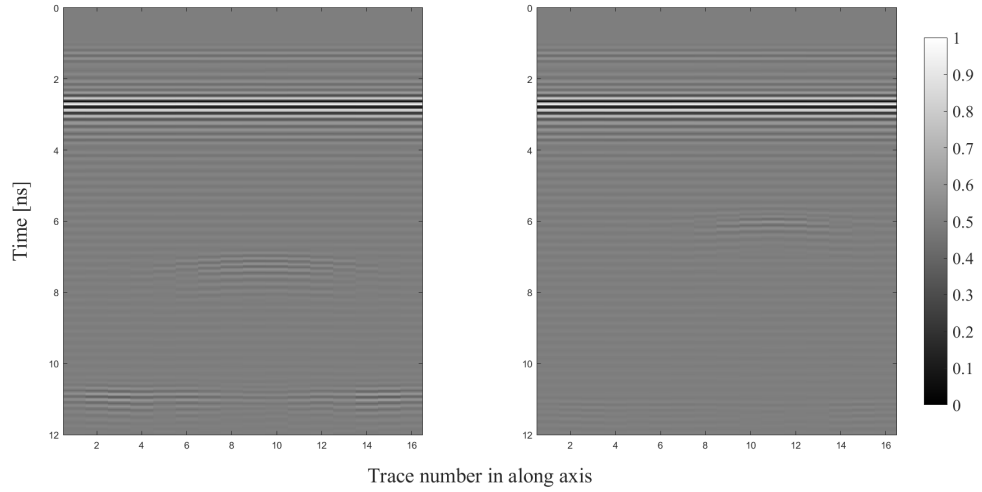
Raw time signals in 2D data (B-scan) are generated as main data set for diversification methods and some A-scan signals and their combined form, B-scan images are demonstrated in Figure 4.3 for sample specific scenarios.

4.3.2 Case 2- Data Set of Extracted Feature Based on Raw Time Signals

Some part of extracted features of data set include PCs of B-scan, they are extracted by using PCA similar as in studies [27, 63]. It is mentioned that B-scan data has the



(a)



(b)

Figure 4.3 Sample of raw A-scan signals and constructed B-scan images from these signals for the following two test scenarios. Raw A-scans for (a) the first scenario corresponds to water content of 2.8 %, depth 306 mm, lateral position 214 mm, and radius 32 mm and the second scenario corresponds to water content of 5.5 %, depth 218 mm, lateral position 244 mm and radius 27 mm (b) B-scan images for the first and the second scenarios

size of 600×16 (600 time steps and 16 A-scans). After applying PCA to raw B-scan, the array of principal components for a B-scan image is obtained as the size of 16×15 . Its rows are the A-scans and the columns are the components. In Figure 4.4, PCs of B-scan for specific scenarios are represented. Some statistical features (SFs) like mean, variance, standard deviation, skewness and kurtosis have been extracted by using following equations (4.1, 4.2, 4.3, 4.4, 4.5). The minimum and maximum features also give information about the reflections from air-ground boundary and features are extracted from unprocessed, raw time signals. For this reason, minimum and

maximum parameters are not evaluated in statistical features. Mean gives the average value of normalized power amplitude (PA) in one A-scan with equation (4.1) where N is 600, discrete-time sample number in 12 ns. By using equation (4.2), variance which is obtained as average amplitude of squared differences from mean value. Another feature is standard deviation which is calculated with equation (4.3) gives the square root of variance, how spread out power amplitude in time-series. Furthermore, skewness which gives the measurement of the asymmetry of the PA around mean value via equation (4.4). The last one is kurtosis calculated with equation (4.5) and it measures the normal distribution of PA, describes shape of tail histogram.

$$M = \frac{1}{N} \sum_{i=1}^N PA_i \quad (4.1)$$

$$V = \frac{1}{N-1} \sum_{i=1}^N |PA_i - M|^2 \quad (4.2)$$

$$STD = \sqrt{\frac{1}{N-1} \sum_{i=1}^N |PA_i - M|^2} \quad (4.3)$$

$$S = \frac{E(PA - M)^3}{STD^3} \quad (4.4)$$

$$K = \frac{E(PA - M)^4}{STD^4} \quad (4.5)$$

where M is the mean, V is the variance, STD is the standard deviation, S is the skewness, K is the kurtosis and $E(x)$ gives the expected value. The extracted features PCs and SFs matrix is transformed to 1D feature vector and its size for each B-scan 1×320 .

4.3.3 Case 3- Data Set of Consecutive Raw A-scans (Time Signals in 1D)

In this case, raw B-scan data matrix is converted to 1D consecutive A-scans and a vector is obtained. Its size is 1×9600 for each B-scan scenario. Figure 4.5 presents some samples of consecutive A-scans for specific scenarios.

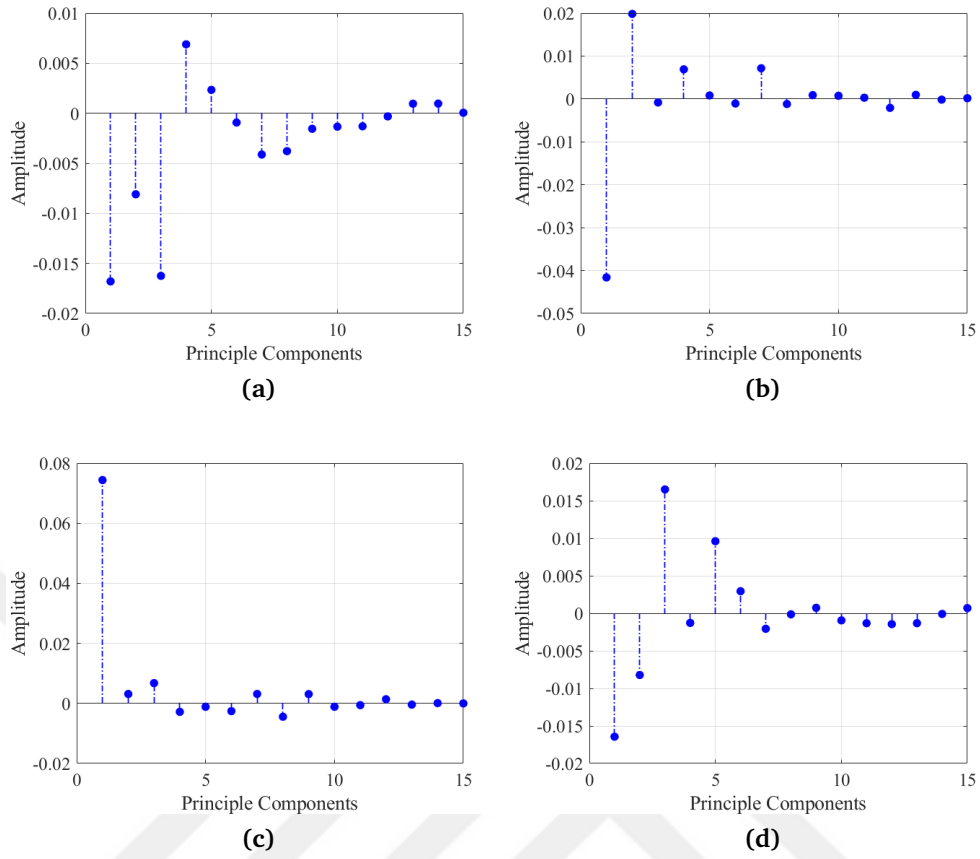


Figure 4.4 Sample principle components of A-scans (A-scan ID: 4, 8, 12 and 16) belonging the scenario corresponds to water content of 5.5%, depth 218 mm, lateral position 244 mm and radius 27 mm; (a) A-scan ID: 4, (b) A-scan ID: 8, (c) A-scan ID: 12 and (d) A-scan ID: 16

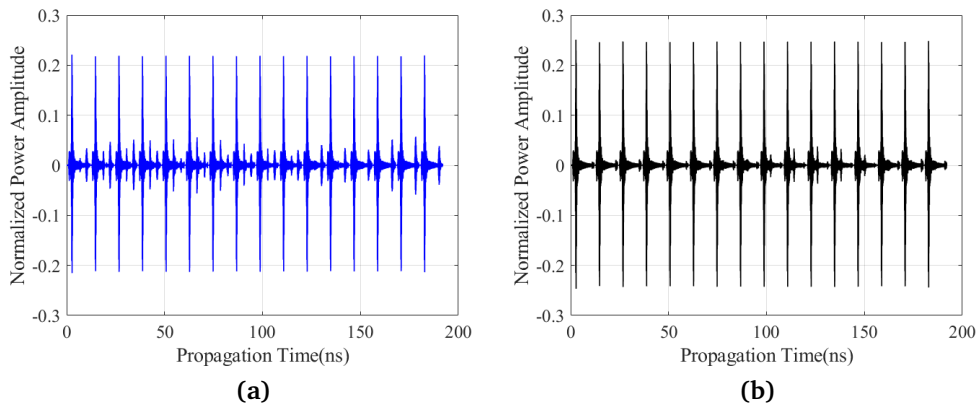


Figure 4.5 Samples of B-scans in a form of consecutive A-scans from test data set for (a) a scenario of water content of 0.2%, depth 311 mm, lateral position 141 mm, and radius 49 mm, (b) a scenario of water content of 5.5%, depth 218 mm, lateral position 244 mm, and radius 27 mm

4.4 Surrogate Modeling of Buried Object Characterization via Deep Regression Network (DRN) Framework Customized for Time Frequency Spectrogram (TFS)

Herein, a novel deep learning framework is introduced as deep regression network DRN which is customized for time frequency spectrogram data version of consecutive A-scans. The proposed framework includes a number of convolutional filter blocks and fully connected (FC) layers [58]. This model uses as input time frequency spectrogram obtained by using STFT of consecutive A-scans. STFT of 1D signals, a sequence of FFT of windowed of 1D signals offer a joint distribution that enables both time and frequency analysis [44]. In addition, the time frequency spectrogram (2D data, image) is utilized for classification of sound signals in some CNN based, deep learning studies [45]. The proposed data type of received reflected power amplitude signal, time frequency spectrogram with DRN framework has been inspired by the previously mentioned developments.

The consecutive A-scans dimension is 1×9600 and window function is selected as 32-point Kaiser window [48] with shape factor 5. Firstly, the signals are zero-averaged. Then, the blocks of the length 32 are extracted over sixteen overlapping intervals and each block is multiplied by the Kaiser window [48] to prevent the spectral leakage to extract the time frequency spectrogram. Subsequently, FFT of the length of 256 are taken. Since the FFT is symmetrical for real signals, the one-sided portion of the image is taken. The magnitude spectrum is calculated to process the TFS as an image. With the conversion of complex numbers to real numbers by taking their magnitude, the TFS data of consecutive A-scan signals are formed and the images are reduced to dimension 128×128 . In Figure 4.6, some TFS images are represented for sample scenarios.

The DRN model consists of seven main blocks as shown in Figure 4.7. The first fourth blocks are used to extract the features from the $128 \times 128 \times 1$ TFS data and down sampling of extracted features. After extracting features of the STFT image in the convolution layers, data processing is followed with the batch normalization (BN) [62] operations and Leaky ReLU layer with the scale parameter of 0.1 (a scalar multiplier of negative inputs) as an activation function. Owing to the global average pooling layer [49] used in the fifth block, a $1 \times 1 \times 256$ feature vector is obtained. Next, sixth block, the feature vector is processed over from 512 neurons to 64 neurons in the FC layer similar with pyramidal shape, and in the last block input is converted to geophysical parameters with regards to depth, lateral position and radius of object via 3 neuron FC layer and regression layer. Figure 4.8 demonstrates the training progress by providing the loss values versus the iteration number.

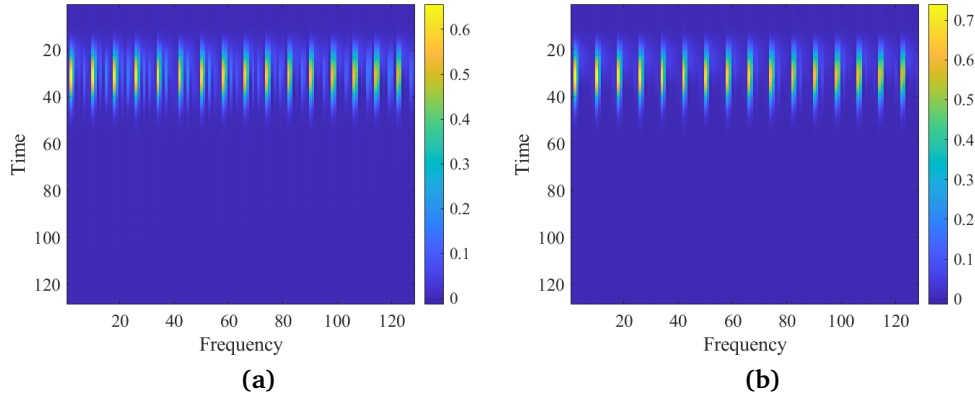


Figure 4.6 Samples of TFS image of consecutive A-scans from test data set for (a) a scenario of water content of 0.2%, depth 311 mm, lateral position 141 mm, and radius 49 mm, (b) a scenario of water content of 5.5%, depth 218 mm, lateral position 244 mm, and radius 27 mm

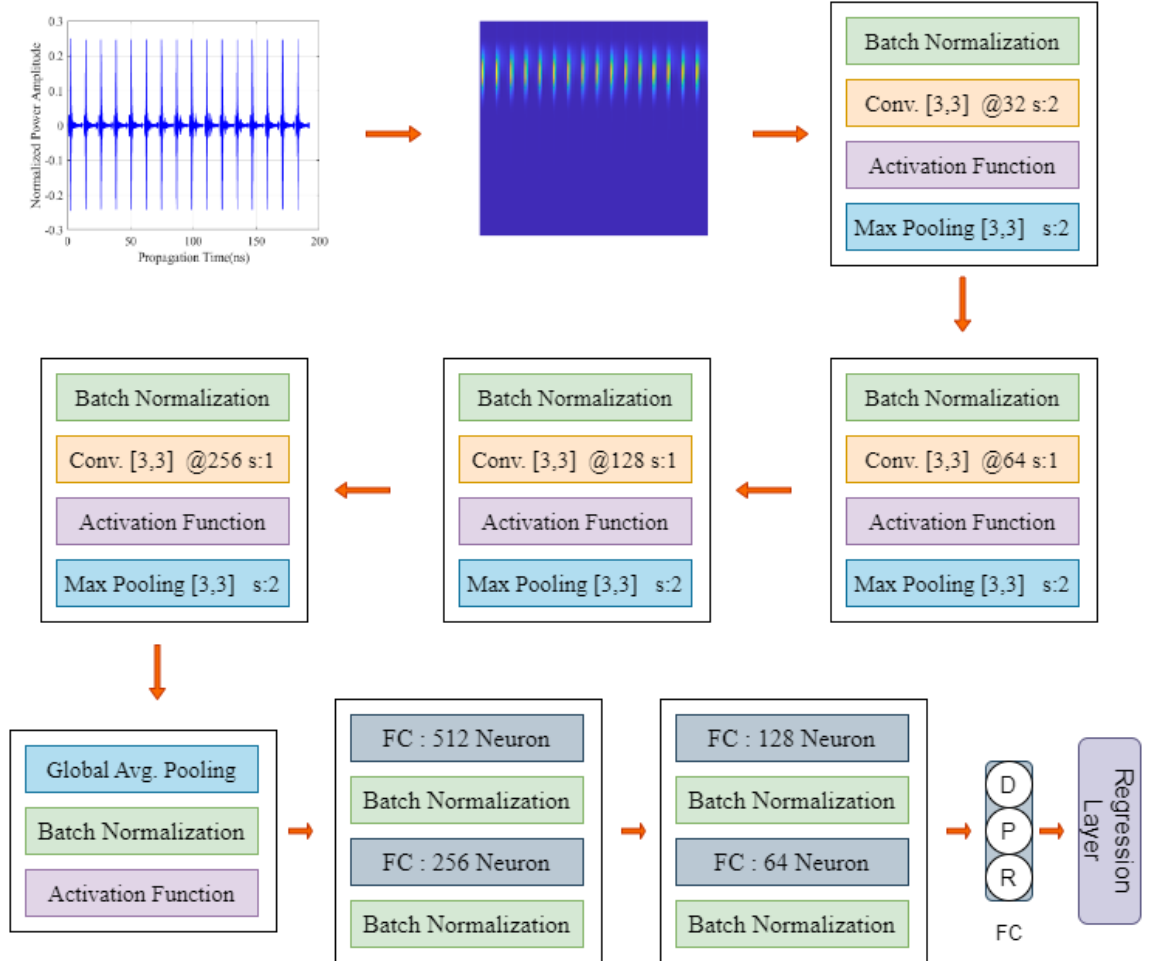


Figure 4.7 The architecture of the proposed deep regression network framework customized for time frequency spectrogram

4.5 Benchmark Models

The mentioned surrogate modelling approach, DRN framework customized for TFS is compared to the benchmark models used for buried object characterization. The benchmark models include CNN [12, 13, 26, 37], MLP [3, 7, 9], GP regression [20], and SVRM [39]. These models are briefly explained.

CNN is a version of deep learning model [10, 24, 25, 28, 38]. The convolutional layer, which is one of the main components of the network, has the ability to automatically extract the data features owing to the convolution filter in the layer. In CNN, several blocks are utilized together, such as a convolution filter, a BN layer, a pooling layer, the activation function, and a fully connected (FC) layer, all involved in the hidden layer. Herein, the architecture and the hyper-parameter configuration of the CNN is used as the following structure which consist of three convolution layers followed by the BN layer, activation function as ReLU layer, the pooling layer included after the last convolution layer and a FC layer with three neurons, and lastly regression layer to estimate the system outputs. Other user defined parameters, such as the size and the number of the convolution filter (32, 64, 128) and pooling layer are assigned according to literature recommendations [13]. Furthermore, it should be emphasized that—in the cited works—the input data is two dimensional, consequently, the filters of the CNN layers are of the corresponding dimensionality, whereas the filters consisted of the CNN model used for benchmarking in this work are arranged according to dimensionality of data sets in the analyzed cases.

The next benchmark model is MLP [3, 7, 9, 11, 42]. The model utilized here features the following hyper-parameter configuration: two hidden layers with 16 and 32 hidden neurons, respectively; log-sigmoid activation functions, training by the Levenberg-Marquardt algorithm until the maximum epoch number reaches 500.

Another surrogate model for benchmarking is GP regression which is an approximation-based machine learning technique widely used for regression and classification approaches. GP regression is based on generalizing Gaussian probability distributions to functions [20, 40]. For this analysis, the kernel function for numerical experiments is selected as “matern3/2” [41], and the prediction method has been defined as block coordinate descent with the block size of 200. The selection of hyper-parameters is an important step in surrogate modelling. Here, Bayesian optimization (an in-built optimization tool in MATLAB) is used for optimum hyper-parameters determination. For validation, the K-fold technique with $K = 10$ has been used.

The last surrogate model is SVRM which belongs to the class of supervised statistical

learning techniques [8, 9, 27, 39]. Herein, SVRM is applied using Bayesian optimization to define hyper-parameters. One of the important components SVRM is a kernel function and it is selected as a Gaussian function for case 3 and Radial Basis Function (RBF) for case 2 to realize nonlinear mapping between the received reflected power amplitudes and geophysical parameters of the buried object.

4.6 The Results of Surrogate Models with Study Cases

Herein, MATLAB is used as coding environment for surrogate modeling of defined inverse GPR problem. The predictive performances of the proposed DRN framework for customized TFS data and all benchmark models are evaluated with the MAE and RME error metrics defined by equations (2.3) and (2.4) respectively. Table 4.3 represents for statistical results over ten independent run for all cases and models with training times and required processing time for a test data. The results show that characterization analysis with case 1- 2D data has the most higher computation cost, after to decrease required time of training the models extracted features are used by the helping of dimensionality reduction methods. However, the prediction performance of the data set of extracted features is much worse than in case 1. For this reason, the characteristic parameters of buried object are estimated with 1D consecutive A-scans (case 3) to realize the aim of computationally efficient characterization together with high prediction performance (low error metrics).

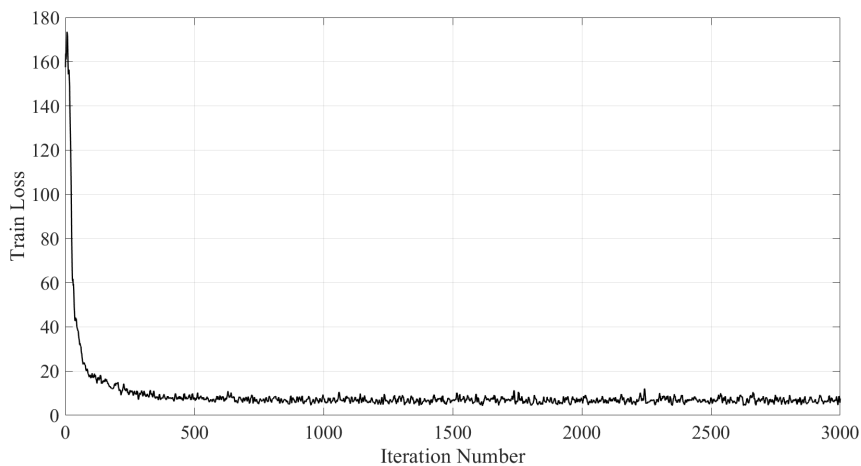


Figure 4.8 The training loss of the proposed DRN model customized for TFS image versus iteration number

In addition, Table 4.3 shows the time-series form of B-scan data in characterization analysis have high computational cost and low prediction performance. However, characterization analysis in case 3 TFS image of consecutive A-scans together with DRN framework provides the most successful prediction performance in a

computationally efficient manner. The proposed data-driven surrogate model is accomplished in solving the defined inverse GPR problem with unprocessed, raw data. Table 4.4 demonstrates the error metrics of individual geophysical parameters obtained from case 3 analysis for the best performances over ten independent runs.

Table 4.3 Prediction performances and computational costs of the all cases and surrogate models, averaged with standard deviation over ten independent runs

Case	Model	MAE [mm]	RME [%]	Training Time [Minutes]	Processing Time for a Single Input Test Data [ms]
Case 1 (2D Raw B-scans)	CNN-2D	12.7±0.8	13.6±1.0	229.0±5.0	18.9±1.2
Case 2 (Feature Based Data)	CNN-2D	17.8±0.9	16.3±0.9	9.0±1.1	7.4±2.2
	CNN-1D	21.2±0.7	19.0±1.3	9.0±1.0	6.5±1.3
	MLP	37.6±2.5	33.8±2.9	29.4±1.2	1.7±0.7
	SVRM	36.8±1.8	43.7±3.5	0.6±0.1	0.4±0.1
	GP-regression	31.4±0.9	31.3±3.4	1.5±0.1	0.3±0.1
Case 3 (1D Consecutive A-scans)	DRN with TFS	3.6±0.2	4.7±0.6	18.3±0.7	7.7±0.6
	CNN-1D	11.8±0.5	11.6±0.6	187.6±6.1	17.6±2.7
	MLP	43.1±3.4	31.4±1.5	360.5±9.1	1.7±0.4
	SVRM	26.1±1.1	24.2±0.8	9.4±0.6	3.5±0.3
	GP-regression	22.9±0.1	21.6±0.2	6.1±0.6	1.1±0.1

It should be mentioned that for case 3 (1×9600 1D consecutive A-scans) MLP model could not be successful in used hardware configuration due to not enough RAM capacity, so down sampling is applied at 10 % and new data dimension is size of 1×960. In case 1, for analysis with SVRM the optimally selected hyper-parameters of Gaussian kernel functions are: Box-constraint of 981.36, Kernel scale of 0.5527 and Epsilon 1.8628. The proposed DRN framework uses the Adam optimizer [61], which is a version of the back propagation algorithm for training the surrogate model. Here, it is used with the batch size of 64 and the maximum epoch number of 500, also TFS data is shuffled in every epoch. Furthermore, another model benchmark model is CNN model, which is trained using same optimizer and batch size with DRN model. Also, the learning rate is set to default value as 10⁻³ until maximum epoch number reached to 500. It is emphasized that the error metrics in Table 4.4 indicate

Table 4.4 The best prediction performance of considered characteristic parameters for case 3 and the proposed model and benchmark models

Model	Characteristic	MAE [mm]	RME [%]	Average MAE [mm]	Average RME [%]
DRN with TFS	Depth	4.2	2.0	3.1	3.9
	Lateral Position	2.8	1.4		
	Radius	2.2	8.3		
CNN-1D	Depth	14.8	6.6	11.4	11.3
	Lateral Position	14.8	7.2		
	Radius	4.7	20.1		
MLP	Depth	47.4	15.7	22.9	21.0
	Lateral Position	14.1	23.4		
	Radius	7.1	32.6		
SVRM	Depth	52.5	26.6	25.1	23.4
	Lateral Position	13.8	6.7		
	Radius	8.9	36.8		
GP- regression	Depth	47.3	23.4	22.8	21.4
	Lateral Position	13.9	6.8		
	Radius	7.3	34.1		

the success of the proposed data TFS with DRN surrogate model in a qualitative sense. In addition, as it can be seen from the results of the benchmark models and the proposed model with TFS version of consecutive A-scans, the performance of prediction of geophysical parameters of buried object independent of each other is satisfactory. Also, it is achieved in a computationally efficient manner even in the case of using limited linear sampled train data set including four different subsurface media. Moreover, Figure 4.9 illustrates the geometrical configurations of some sample test scenarios in terms of alignment between predicted with proposed DRN surrogate model and target parameters of the object. Table 4.5 represents numerical data, i.e., the target and predicted by surrogate models characteristic parameters for some test scenarios. As it can be observed from Figure 4.9, agreement in the geometrical representations between the actual and predicted object size and location is excellent. Lastly, the results of benchmarking analysis are demonstrated with Table 4.6 as numerical and Figure 4.10 as geometrical configuration in terms of the target and predicted parameters for case 3.

The error metrics of the DRN framework is 3.1 mm (MAE) and 3.9 % (RME), which is the average of considered characteristic parameters (depth, lateral position and radius). The second best benchmark model, CNN exhibits considerably worse accuracy with the average error (11.4 mm MAE and 11.3 % RME) being almost three times as high.

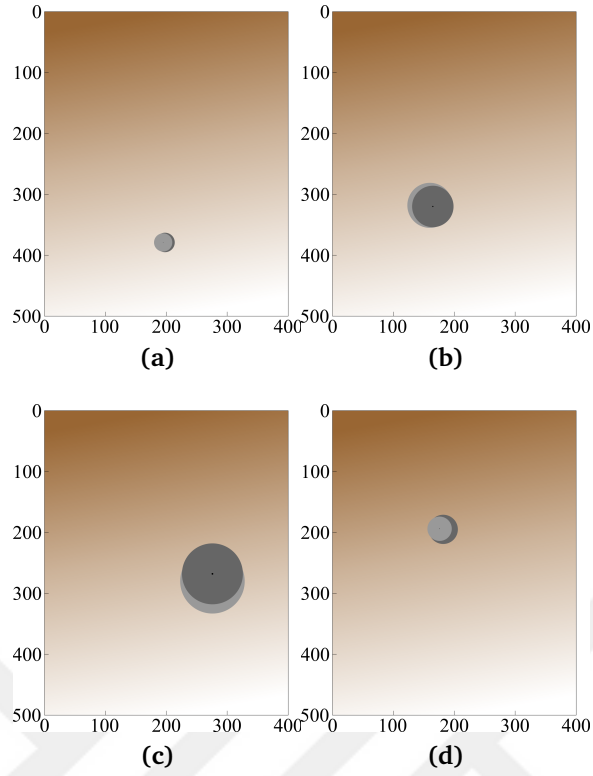


Figure 4.9 Prediction of characteristic parameters of the object, obtained using the proposed surrogate model DRN. The target and surrogate-predicted object marked using the dark- and light-grey shade in the geometrical configuration, respectively: (a) $D = 379$ mm, $P = 198$ mm, $R = 16$ mm, (b) $D = 320$ mm, $P = 165$ mm, $R = 34$ mm, (c) $D = 268$ mm, $P = 276$ mm, $R = 50$ mm, (d) $D = 195$ mm, $P = 182$ mm, $R = 24$ mm

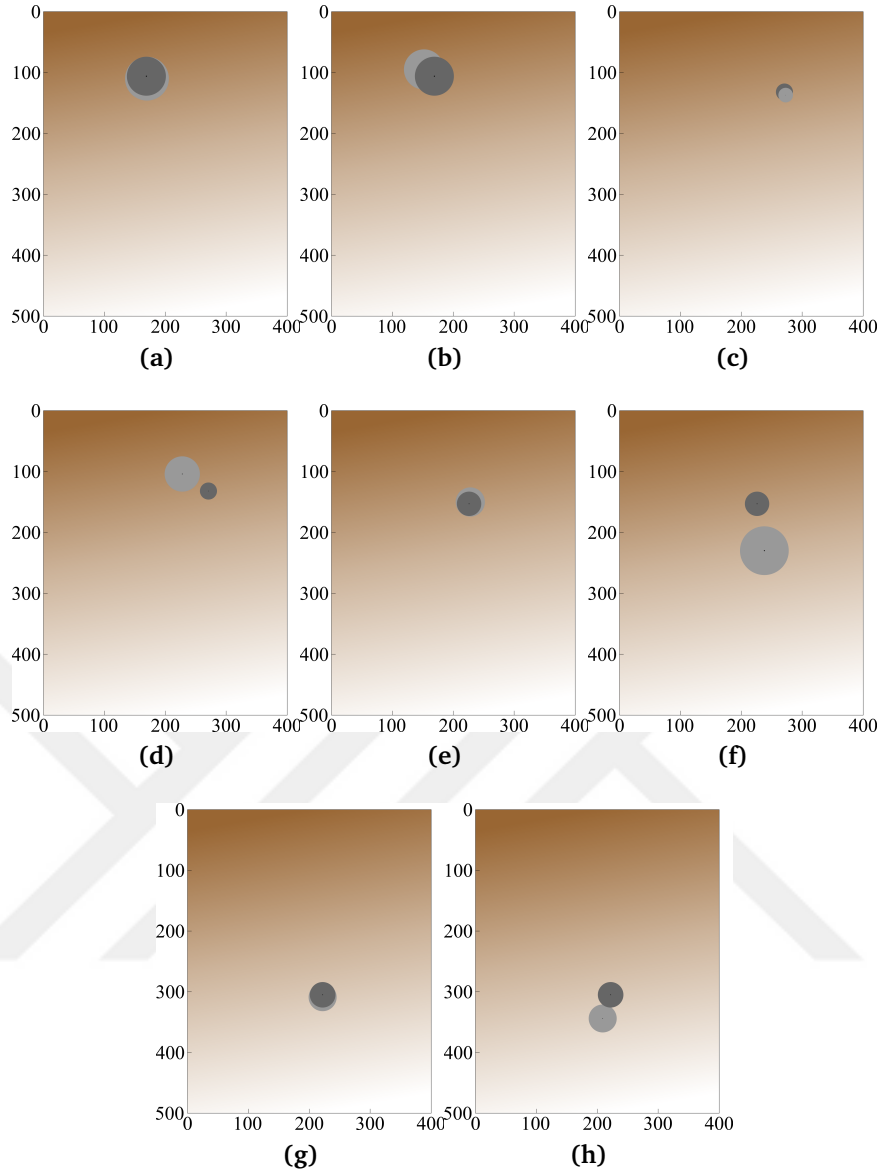


Figure 4.10 Prediction of characteristic parameters of the object, obtained using the proposed surrogate model DRN and the benchmark models. The target and surrogate-predicted object marked using the dark- and light-grey shade in the geometrical configuration, respectively: Scenario I: $D = 193$ mm, $P = 110$ mm, $R = 34$ mm, (a) proposed model, (b) CNN; Scenario II: $D = 361$ mm, $P = 189$ mm, $R = 21$ mm, (c) proposed model, (d) MLP; Scenario III: $D = 244$ mm, $P = 299$ mm, $R = 26$ mm, (e) proposed model, (f) SVRM; Scenario IV: $D = 193$ mm, $P = 110$ mm, $R = 34$ mm, (g) proposed model, (h) GP-regression

Table 4.5 DRN framework customized for TFS image estimation of the characteristic parameters in comparison with their true values for selected test scenarios

Model	<i>D</i>	<i>P</i>	<i>R</i>	Error (<i>D, P, R</i>) [mm]
True Value	290	144	15	-
DRN with TFS	288	144	13	2, 0, 2
True Value	352	231	22	-
DRN with TFS	353	230	24	-1, 1, -2
True Value	274	170	34	-
DRN with TFS	274	166	35	0, 4, -1
True Value	287	142	31	-
DRN with TFS	287	147	31	0, -5, 0

Table 4.6 Characteristic parameter prediction for the proposed DRN and benchmark methods in comparison with their true values for selected test scenarios

Model	<i>D</i>	<i>P</i>	<i>R</i>	Error (<i>D, P, R</i>) [mm]
True Value	155	224	38	-
DRN with TFS	146	224	40	9, 0, -2
CNN-1D	176	260	39	-21, -36, -1
MLP	309	238	34	-154, -14, 4
SVRM	180	228	46	-25, -4, -8
GP- regression	184	220	39	-29, 4, -1
True Value	110	178	32	-
DRN with TFS	113	175	30	-3, 3, 2
CNN-1D	105	147	26	5, 31, 6
MLP	103	229	29	7, -51, 3
SVRM	164	187	40	-54, -9, -8
GP- regression	159	198	38	-49, -20, -6
True Value	211	225	15	-
DRN with TFS	209	232	17	2, -7, -2
CNN-1D	234	217	25	-23, 8, -10
MLP	367	185	29	-156, 40, -14
SVRM	363	209	21	-152, 16, -6
GP- regression	339	210	30	-128, 15, -15

4.7 Characterization Analysis for Validation of Case 3 Under Realistic Scenarios with Noisy Data

Herein, for the purpose of validation of the proposed surrogate model and data set, new data sets are obtained by random noise addition [14, 17, 25, 32–36] to the generated raw consecutive A-scans. In the literature, different approaches to noise incorporation are offered as mentioned in Chapter 2 and 3. The particular problem considered in this chapter is solved by using the TFS image, consecutive 1D time-varying amplitude signals (A-scans), and noise analysis is studied for case 3 and similar purpose in [14, 17] which is to bring the models closer to the real-time

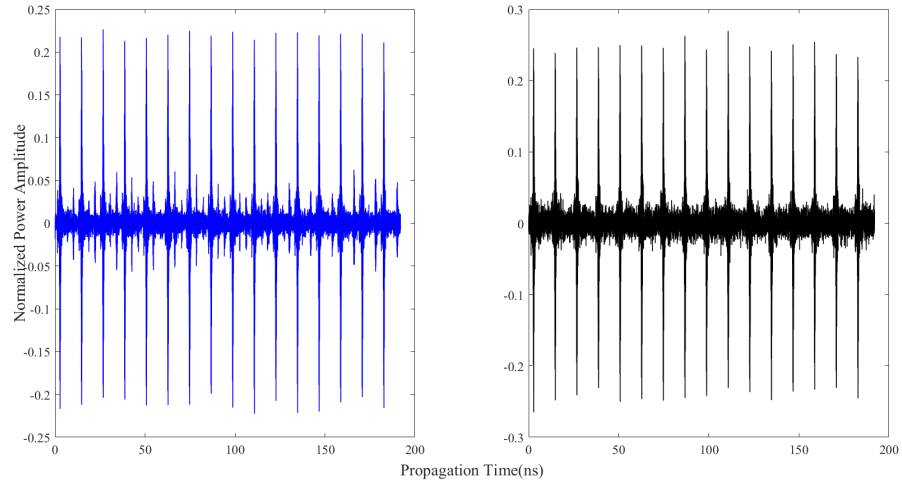
applications. Hence, for further verification of DRN framework customized for TFS, supplementary noisy data sets are generated with two different SNR levels as 20 dB and 30 dB by adding white Gaussian noise [25]. In Figure 4.11, noisy B-scans and their converted version of 1D consecutive A-scans are presented with two different scenarios and different SNR levels of 20 dB and 30 dB. For this analysis, the proposed DRN framework is compared with CNN benchmark model. The results are presented in Table 4.7 and Table 4.8 provides a comparison of model predicted versus actual characteristic parameters for selected test cases. The results indicate that the DRN framework outperforms the benchmark by a considerable margin.

Table 4.7 Prediction performances of characteristic parameters of the buried object for DRN and CNN models. Considered scenarios include noisy data with SNR of 30 dB and 20 dB

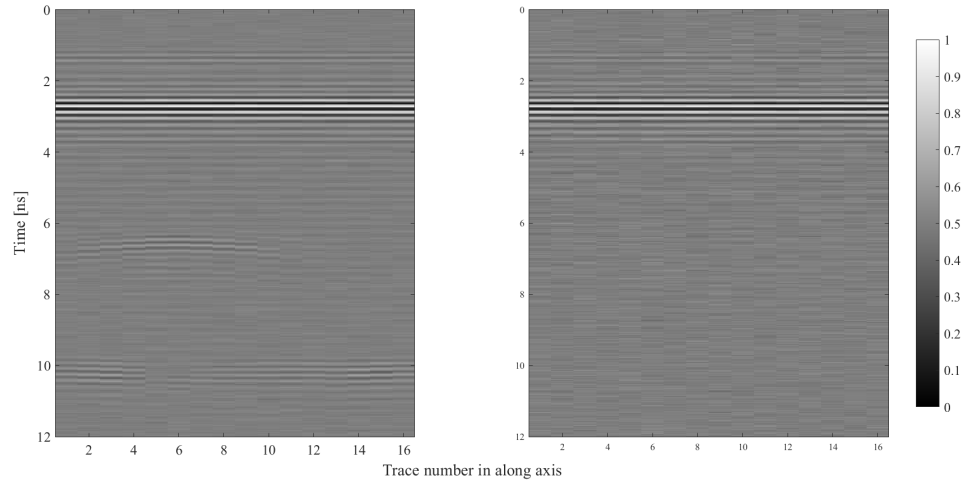
Model	Characteristic	SNR = 30dB			SNR = 20dB		
		MAE [mm]	RME [%]	Average MAE [mm]	MAE [mm]	RME [%]	Average MAE [mm]
DRN	Depth	13.1	5.4	10.1	21.6	9.4	15.8
	Lateral Position	10.5	5.3		15.2	7.7	
	Radius	6.5	26.0		6.8	30.1	
CNN	Depth	31.5	12.4	20.7	40.7	16.4	25.2
	Lateral Position	23.6	10.7		27.6	12.6	
	Radius	6.9	29.5		7.4	32.0	

Table 4.8 Characteristic parameter prediction for the DRN and CNN models in comparison with their true values for selected noisy test scenarios

Model	<i>D</i>	<i>P</i>	<i>R</i>	Error (<i>D</i> , <i>P</i> , <i>R</i>) [mm]
True Value	128	160	19	-
DRN [30 dB]	132	164	25	-4, -4, -6
CNN [30 dB]	171	145	30	-43, 15, -11
True Value	240	150	14	-
DRN [20 dB]	247	158	18	-7, -8, -4
CNN [20 dB]	255	169	16	-15, -19, -2
True Value	160	232	45	-
DRN [30 dB]	167	232	42	-7, 0, 3
CNN [30 dB]	130	215	33	30, 17, 12
True Value	276	201	21	-
DRN [20 dB]	281	216	24	-5, -15, -3
CNN [20 dB]	263	178	27	13, 23, -6



(a)



(b)

Figure 4.11 Samples of noisy B-scans in a form of consecutive A-scans from test data set for the first scenario SNR value of 30 dB scenario of water content of 0.2%, depth 311 mm, lateral position 141 mm, and radius 49 mm; for the second scenario SNR value of 20 dB scenario of water content of 6.2%, depth 211 mm, lateral position 225 mm, and radius 15 mm (a) consecutive A-scans and (b) B-scan images for the first and the second scenarios

5

RESULTS AND DISCUSSION

In this thesis, artificial-intelligence-based surrogate modeling is proposed for buried object characterization with regards to depth, lateral position and radius. Herein, the goal of this study propose computationally efficient, novel deep learning-based frameworks to solve defined problem of parametric estimation of characteristic parameters of buried object. With this motivation, novel surrogate modeling approaches using different user-defined data sets are proposed. The presented first approach is a deep-learning-based TFRM framework which has the ability of constructing accurate representation of the buried object characteristics using sparse training data set. This surrogate model has been comprehensively validated using a number of specific cases of perfectly electric conductor (PEC)-object buried in different, dispersive soil media, at various positions. The results indicate competitive performance of the proposed technique with the MAE value of less than 12 (as compared to MAE of 23 for the best benchmark approach, CNN). Another approach is utilization of the hyperbolic signatures, extracted from the pre-processed B-scans by using linear regression technique, as the inputs of the modeling process. The proposed novel regression-assisted surrogate model is M2LP framework consisting of deep-learning-based feedforward network configuration in the form of a series of blocks of fully connected layers, batch normalization, and activation layers for this approach. The results indicate competitive prediction performance of M2LP with the average MAE as low as 10 mm versus 18.3 mm for the best benchmark model, CNN.

In the Chapter 4, a comprehensive analysis is introduced for buried object characterization via surrogate models and some cases such as using raw, unprocessed 2D B-scan data, extracted features from raw A-scans and a new 1D consecutive raw A-scans data. The results indicate that modeling with B-scan has the highest computational cost and feature extraction methodology decreases the required training time but prediction performance is getting worse. Also, the case study of consecutive A-scans with conventional CNN models not only takes so much time for training but also the prediction performance is not applicable to estimate characteristic

parameters of buried object as an inverse GPR problem. Hence, a novel surrogate model is presented as fast and accurate DRN framework customized for TFS. The performance results which is the average of considered characteristic parameters (depth, lateral position and radius) show competitive prediction performance of DRN with the average 3.1 mm (MAE) and 3.9 % (RME), versus 11.4 mm (MAE) and 11.3 % (RME) for the best benchmark model, CNN.

The approaches presented in the thesis can be considered a viable alternative to existing surrogate modeling techniques for buried object characterization both with respect to high accuracy rate and computational efficiency. On the other hand, the proposed methodologies present some limitations, in particular, characterization objects made of different material types, shapes and buried multiple objects. It should be emphasized that the presented models are expandable and improvable. One of possible options, to be considered as a part of the future work, is to enable identification of the buried objects different material types and shapes by adding the mentioned data structures as extra inputs of the underlying surrogate model.

REFERENCES

- [1] D. J. Daniels, "System design," in *Ground Penetrating Radar*, London, United Kingdom: Institution of Engineering and Technology, 2004, pp. 13–36.
- [2] H. M. Jol, "Electromagnetic principles of ground penetrating radar," in *Ground Penetrating Radar Theory and Applications*, Amsterdam: Elsevier, 2009, pp. 5–17.
- [3] T. Liu, Y. Su, C. Huang, "Inversion of ground penetrating radar data based on neural networks," *Remote Sensing*, vol. 10, no. 5, 2018. DOI: 10.3390/rs10050730.
- [4] C. Özdemir, Ş. Demirci, E. Yiğit, B. Yilmaz, "A review on migration methods in b-scan ground penetrating radar imaging," *Mathematical Problems in Engineering*, vol. 2014, pp. 1–16, 2014.
- [5] P. Sharma, B. Kumar, D. Singh, S. Gaba, "Critical analysis of background subtraction techniques on real GPR data.," *Defence Science Journal*, vol. 67, no. 5, pp. 559–571, 2017.
- [6] H. Liu, C. Lin, J. Cui, L. Fan, X. Xie, B. F. Spencer, "Detection and localization of rebar in concrete by deep learning using ground penetrating radar," *Automation in Construction*, vol. 118, p. 103279, 2020.
- [7] Q. Dou, L. Wei, D. R. Magee, A. G. Cohn, "Real-time hyperbola recognition and fitting in GPR data," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 55, no. 1, pp. 51–62, 2016.
- [8] E. Pasolli, F. Melgani, M. Donelli, "Automatic analysis of GPR images: A pattern-recognition approach," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 47, no. 7, pp. 2206–2217, 2009.
- [9] Y. Jin, Y. Duan, "Wavelet scattering network-based machine learning for ground penetrating radar imaging: Application in pipeline identification," *Remote Sensing*, vol. 12, no. 21, p. 3655, 2020.
- [10] C. Ozdemir, S. Demirci, E. Yigit, A. Kavak, "A hyperbolic summation method to focus B-scan ground penetrating radar images: An experimental study with a stepped frequency system," *Microwave and Optical Technology Letters*, vol. 49, no. 3, pp. 671–676, 2007.
- [11] Y. Zhang, D. Huston, T. Xia, "Underground object characterization based on neural networks for ground penetrating radar data," in *Nondestructive characterization and monitoring of advanced materials, aerospace, and civil infrastructure 2016*, SPIE, vol. 9804, Apr. 2016, pp. 10–18.

- [12] R. T. Sakaguchi, K. D. M. Jr., L. M. Collins, P. A. Torrione, "Recognizing subsurface target responses in ground penetrating radar data using convolutional neural networks," in *Detection and Sensing of Mines, Explosive Objects, and Obscured Targets XX*, S. S. Bishop, J. C. Isaacs, Eds., International Society for Optics and Photonics, vol. 9454, SPIE, 2015, 94541A. DOI: 10 . 1117/12.2177747.
- [13] M. Moalla, H. Frigui, A. Karem, A. Bouzid, "Application of convolutional and recurrent neural networks for buried threat detection using ground penetrating radar data," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 58, no. 10, pp. 7022–7034, 2020.
- [14] W. Lei, J. Luo, F. Hou, L. Xu, R. Wang, X. Jiang, "Underground cylindrical objects detection and diameter identification in GPR B-scans via the CNN-LSTM framework," *Electronics*, vol. 9, no. 11, p. 1804, 2020.
- [15] D. Kumlu, I. Erer, "The multiscale directional neighborhood filter and its application to clutter removal in GPR data," *Signal, Image and Video Processing*, vol. 12, pp. 1237–1244, 2018.
- [16] H. Wang, S. Ouyang, Q. Liu, K. Liao, L. Zhou, "Buried target detection method for ground penetrating radar based on deep learning," *Journal of Applied Remote Sensing*, vol. 16, no. 1, p. 018 503, 2022.
- [17] W. Lei *et al.*, "Automatic hyperbola detection and fitting in GPR B-scan image," *Automation in Construction*, vol. 106, p. 102 839, 2019.
- [18] R. Ahmadi, N. Fathianpour, "Estimating geometrical parameters of cylindrical targets detected by ground-penetrating radar using template matching algorithm," *Arabian Journal of Geosciences*, vol. 10, no. 6, pp. 140–152, 2017.
- [19] X. Zhang, L. Han, M. Robinson, A. Gallagher, "A GANs-based deep learning framework for automatic subsurface object recognition from ground penetrating radar data," *IEEE Access*, vol. 9, pp. 39 009–39 018, 2021.
- [20] E. Pasolli, F. Melgani, M. Donelli, "Gaussian process approach to buried object size estimation in GPR images," *IEEE Geoscience and Remote Sensing Letters*, vol. 7, no. 1, pp. 141–145, 2009.
- [21] U. Ozkaya, F. Melgani, M. B. Bejiga, L. Seyfi, M. Donelli, "Gpr B scan image analysis with deep learning methods," *Measurement*, vol. 165, p. 107 770, 2020.
- [22] A. Giannopoulos, "Modelling ground penetrating radar by GprMax," *Construction and building materials*, vol. 19, no. 10, pp. 755–762, 2005.
- [23] C. Warren, A. Giannopoulos, I. Giannakis, "gprMax: Open source software to simulate electromagnetic wave propagation for ground penetrating radar," *Computer Physics Communications*, vol. 209, pp. 163–170, 2016.
- [24] "GPRInvNet: Deep learning-based ground-penetrating radar data inversion for tunnel linings," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 59, no. 10, pp. 8305–8325, 2021.
- [25] J. Wang, H. Liu, P. Jiang, Z. Wang, Q. Sui, F. Zhang, "GPRI2Net: A deep-neural-network-based ground penetrating radar data inversion and object identification framework for consecutive and long survey lines," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 60, pp. 1–20, 2022.

- [26] Y. Ji *et al.*, “Deep neural network-based permittivity inversions for ground penetrating radar data,” *IEEE Sensors Journal*, vol. 21, no. 6, pp. 8172–8183, 2021.
- [27] N. Smitha, V. Singh, “Target detection using supervised machine learning algorithms for gpr data,” *Sensing and Imaging*, vol. 21, no. 1, p. 11, 2020.
- [28] I. Giannakis, A. Giannopoulos, C. Warren, “A machine learning scheme for estimating the diameter of reinforcing bars using ground penetrating radar,” *IEEE Geoscience and Remote Sensing Letters*, vol. 18, no. 3, pp. 461–465, 2020.
- [29] I. Giannakis, A. Giannopoulos, C. Warren, “A machine learning-based fast-forward solver for ground penetrating radar with application to full-waveform inversion,” *IEEE Transactions on Geoscience and Remote Sensing*, vol. 57, no. 7, pp. 4417–4426, 2019.
- [30] T. Bourdi, J. E. Rhazi, F. Boone, G. Ballivy, “Modelling dielectric-constant values of concrete: An aid to shielding effectiveness prediction and ground-penetrating radar wave technique interpretation,” *Journal of Physics D: Applied Physics*, vol. 45, no. 40, p. 405 401, 2012.
- [31] J. C. Helton, F. J. Davis, “Latin hypercube sampling and the propagation of uncertainty in analyses of complex systems,” *Reliability Engineering & System Safety*, vol. 81, no. 1, pp. 23–69, 2003.
- [32] A. Kumar, U. K. Singh, B. Pradhan, “Ground penetrating radar in coastal hazard mitigation studies using deep convolutional neural networks,” *Remote Sensing*, vol. 14, no. 19, p. 4899, 2022.
- [33] D. Feng, S. Liu, J. Yang, X. Wang, X. Wang, “The noise attenuation and stochastic clutter removal of ground penetrating radar based on the K-SVD dictionary learning,” *IEEE Access*, vol. 9, pp. 74 879–74 890, 2021.
- [34] W. Xue, X. Dai, J. Zhu, Y. Luo, Y. Yang, “A noise suppression method of ground penetrating radar based on eemd and permutation entropy,” *IEEE Geoscience and Remote Sensing Letters*, vol. 16, no. 10, pp. 1625–1629, 2019.
- [35] R. Li *et al.*, “Denoising method of ground-penetrating radar signal based on independent component analysis with multifractal spectrum,” *Measurement*, vol. 192, p. 110 886, 2022.
- [36] H. Hu *et al.*, “A study of automatic recognition and localization of pipeline for ground penetrating radar based on deep learning,” *IEEE Geoscience and Remote Sensing Letters*, vol. 19, pp. 1–5, 2022.
- [37] N. Barkataki, B. Tiru, U. Sarma, “A CNN model for predicting size of buried objects from gpr b-scans,” *Journal of Applied Geophysics*, vol. 200, p. 104 620, 2022.
- [38] R. Yurt, H. Torpi, P. Mahouti, A. Kızılay, S. Koziel, *Buried object characterization using ground penetrating radar*, 2022. DOI: 10.21227/yskj-v875. [Online]. Available: <https://dx.doi.org/10.21227/yskj-v875>.
- [39] S. Hosseinzadeh, M. Shaghaghi, “GPR data regression and clustering by the fuzzy support vector machine and regression,” *Progress In Electromagnetics Research M*, vol. 93, pp. 175–184, 2020.

- [40] J. P. Jacobs, S. Koziel, "Two-stage framework for efficient gaussian process modeling of antenna input characteristics," *IEEE transactions on antennas and propagation*, vol. 62, no. 2, pp. 706–713, 2013.
- [41] A. Melkumyan, F. Ramos, "Multi-kernel gaussian processes," in *Twenty-second international joint conference on artificial intelligence*, 2011.
- [42] P. Mahouti, "Design optimization of a pattern reconfigurable microstrip antenna using differential evolution and 3D EM simulation-based neural network model," *International Journal of RF and Microwave Computer-Aided Engineering*, vol. 29, no. 8, e21796, 2019.
- [43] P. Kim, *Matlab deep learning with machine learning, neural networks and artificial intelligence*. Springer, 2017.
- [44] J. B. Allen, L. R. Rabiner, "A unified approach to short-time Fourier analysis and synthesis," *Proceedings of the IEEE*, vol. 65, no. 11, pp. 1558–1564, 1977.
- [45] Z. Wang, L. Cao, Z. Zhang, X. Gong, Y. Sun, H. Wang, "Short time fourier transformation and deep neural networks for motor imagery brain computer interface recognition," *Concurrency and Computation: Practice and Experience*, vol. 30, no. 23, e4413, 2018.
- [46] A. Anaby-Tavor *et al.*, "Do not have enough data? deep learning to the rescue!" In *Proceedings of the AAAI Conference on Artificial Intelligence*, vol. 34, 2020, pp. 7383–7390.
- [47] S. Koziel, N. Çalık, P. Mahouti, M. A. Belen, "Accurate modeling of antenna structures by means of domain confinement and pyramidal deep neural networks," *IEEE Transactions on Antennas and Propagation*, vol. 70, no. 3, pp. 2174–2188, 2021.
- [48] Y.-P. Lin, P. Vaidyanathan, "A kaiser window approach for the design of prototype filters of cosine modulated filterbanks," *IEEE signal processing letters*, vol. 5, no. 6, pp. 132–134, 1998.
- [49] T.-Y. Hsiao, Y.-C. Chang, C.-T. Chiu, "Filter-based deep-compression with global average pooling for convolutional networks," in *2018 IEEE International Workshop on Signal Processing Systems (SiPS)*, IEEE, 2018, pp. 247–251.
- [50] A. Turk, "UWB performance analysis of PDTEM horn antenna designed for multi-sensor adaptive hand-held GPR," in *11th International Symposium on Antenna Technology and Applied Electromagnetics [ANTEM 2005]*, IEEE, 2005, pp. 1–4.
- [51] A. S. Turk, "Ultra-wideband tem horn design for ground penetrating impulse radar systems," *Microwave and optical technology letters*, vol. 41, no. 5, pp. 333–336, 2004.
- [52] A. S. Turk, S. Aksoy, A. K. Keskin, M. D. Senturk, A. Caliskan, M. B. Ozakin, "Ultra-wide band antenna designs and numerical system modelling for forward-looking GPR," *Near Surface Geophysics*, vol. 13, no. 3, pp. 261–267, 2015.
- [53] A. S. TÜRK, A. K. Keskin, M. D. ŞENTÜRK, "Dielectric loaded TEM horn-fed ridged horn antenna designfor ultrawideband ground-penetrating impulse radar," *Turkish Journal of Electrical Engineering and Computer Sciences*, vol. 23, no. 5, pp. 1479–1488, 2015.

- [54] T.-T. Wong, P.-Y. Yeh, "Reliable accuracy estimates from k-fold cross validation," *IEEE Transactions on Knowledge and Data Engineering*, vol. 32, no. 8, pp. 1586–1594, 2019.
- [55] S. Chapra, "General linear least-squares and nonlinear regression," in *Applied Numerical Methods with MATLAB for Engineers and Scientists*, New York, USA: McGraw-Hill Education, 2018, pp. 385–389.
- [56] Y. Jaluria, "Numerical curve fitting and interpolation," in *Computer methods for engineering with MATLAB applications*, New York, USA: Taylor & Francis, 2011, pp. 252–257.
- [57] R. Yurt, H. Torpi, "Ground penetrating radar data analysis with nonlinear regression on artificial neural network," in *2020 International Congress on Human-Computer Interaction, Optimization and Robotic Applications (HORA)*, 2020, pp. 1–5.
- [58] N. Calik, M. A. Belen, P. Mahouti, "Deep learning base modified mlp model for precise scattering parameter prediction of capacitive feed antenna," *International journal of numerical modelling: electronic networks, devices and fields*, vol. 33, no. 2, e2682, 2020.
- [59] E. Temlioglu, I. Erer, "A novel convolutional autoencoder-based clutter removal method for buried threat detection in ground-penetrating radar," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 60, pp. 1–13, 2021.
- [60] I. Giannakis, F. Zhou, C. Warren, A. Giannopoulos, "On the limitations of hyperbola fitting for estimating the radius of cylindrical targets in nondestructive testing and utility detection," *IEEE Geoscience and Remote Sensing Letters*, vol. 19, pp. 1–5, 2022.
- [61] D. P. Kingma, J. Ba, "Adam: A method for stochastic optimization," *arXiv preprint arXiv:1412.6980*, 2014.
- [62] S. Ioffe, C. Szegedy, "Batch normalization: Accelerating deep network training by reducing internal covariate shift," in *International conference on machine learning*, pmlr, 2015, pp. 448–456.
- [63] B. Yoldemir, M. Sezgin, "Peak scatter-based buried object identification using GPR-EMI dual sensor system," *Nondestructive Testing and Evaluation*, 2019.
- [64] J. Zhu, W. Xue, X. Rong, Y. Yu, "A clutter suppression method based on improved principal component selection rule for ground penetrating radar," *Progress In Electromagnetics Research M*, vol. 53, pp. 29–39, 2017.
- [65] M. Abd Shehab, M. A. S. Al Obaidi, I. Hos, S. Karamzadeh, "Subspace clutter removal techniques in gpr images," *Progress In Electromagnetics Research M*, vol. 82, pp. 139–147, 2019.
- [66] R. Yurt, H. Torpi, P. Mahouti, A. Kizilay, S. Koziel, "Buried object characterization using ground penetrating radar assisted by data-driven surrogate-models," *IEEE Access*, vol. 11, pp. 13 309–13 323, 2023. DOI: 10 . 1109/ACCESS.2023.3243132.
- [67] P. K. Verma, A. N. Gaikwad, D. Singh, M. Nigam, "Analysis of clutter reduction techniques for through wall imaging in UWB range," *Progress In Electromagnetics Research B*, vol. 17, pp. 29–48, 2009.

- [68] E. Temlioglu, I. Erer, "Clutter removal in ground-penetrating radar images using morphological component analysis," *IEEE Geoscience and Remote Sensing Letters*, vol. 13, no. 12, pp. 1802–1806, 2016.
- [69] Q. Lu, J. Pu, Z. Liu, "Feature extraction and automatic material classification of underground objects from ground penetrating radar data," *Journal of Electrical and Computer Engineering*, vol. 2014, pp. 28–28, 2014.
- [70] M. S. El-Mahallawy, M. Hashim, "Material classification of underground utilities from gpr images using DCT-based SVM approach," *IEEE Geoscience and Remote Sensing Letters*, vol. 10, no. 6, pp. 1542–1546, 2013.
- [71] F. Queiroz, D. Vieira, X. Travassos, "Analyzing the relevant features of GPR scattered waves in time-and frequency-domain," *Research in Nondestructive Evaluation*, vol. 24, no. 2, pp. 105–123, 2013.



PUBLICATIONS FROM THE THESIS

Papers

1. R. Yurt, H. Torpi, P. Mahouti, A. Kizilay and S. Koziel, "Buried Object Characterization Using Ground Penetrating Radar Assisted by Data-Driven Surrogate-Models," in *IEEE Access*, vol. 11, pp. 13309-13323, 2023. doi: 10.1109/ACCESS.2023.3243132.
2. R. Yurt, H. Torpi, A. Kizilay, S. Koziel, A. Pietrenko-Dabrowska, and P. Mahouti, "Buried object characterization by data-driven surrogates and regression-enabled hyperbolic signature extraction," *Scientific Reports*, vol. 13, no. 1, pp. 5717-5739, 2023, doi: 10.1038/s41598-023-32925-6.

Conference Papers

1. R. Yurt and H. Torpi, "Ground Penetrating Radar Data Analysis with Nonlinear Regression on Artificial Neural Network," *2020 International Congress on Human-Computer Interaction, Optimization and Robotic Applications (HORA)*, Ankara, Turkey, 2020, pp. 1-5, doi: 10.1109/HORA49412.2020.9152599.

Projects

1. "Study and development of target imaging and recognition technologies for contact and forward-looking GPR systems", supported by The Scientific and Technological Research Council of Turkey. Project No: 119N196.