Journal of Physics:
Conference Series

PURPOSE-LED PUBLISHING™

PAPER • OPEN ACCESS

Classification Size of Underground Object from Ground Penetrating Radar Image using Machine Learning Technique

To cite this article: Mohd Shuhanaz Zanar Azalan *et al* 2023 *J. Phys.: Conf. Ser.* **2550** 012029

Classification Size of Underground Object from Ground Penetrating Radar Image using Machine Learning Technique

Mohd Shuhanaz Zanar Azalan*[1], Tang Esian[1], Hasimah Ali[1], Ahmad Firdaus Ahmad Zaidi[1], Tengku Sarah Tengku Amran[2]

[1]Faculty of Electric Engineering Technology UniMAP Arau Perlis
[2]Nuclear Agency Malaysia, Bangi, Selangor

Email: shuhanaz@unimap.edu.my

Abstract. Ground Penetrating Radar (GPR) is a useful tool in detecting subsurface object or hidden structure defects However, the time-consuming problems and high requirement of professional manpower is required to analyse the GPR data. Machine learning is a tool that endowed with the ability to learn, and it can reduce time taken for the GPR data analysing. To simplify the identification process, a framework is proposed to classify the size of underground metallic pipe by using Histogram of Oriented Gradient (HOG) as a feature extraction algorithm. Two machine learning algorithms namely Support Vector Machines (SVM) and Backpropagation Neural Network were proposed to classify the size of the underground metallic pipe. As a result, the accuracy from the identification is more than 98% for both classifier algorithm.

Keywords: Ground Penetrating Radar, Machine Learning, Underground Object Size

1. Introduction

Ground Penetrating Radar (GPR) is a high-resolution subsurface object detection system. GPR can sends a high-frequency electromagnetic pulse penetrate the soil at a speed determined by the material's dielectric things. GPR helps to detect what beneath underground or in the buildings. It is widely used to avoid potential threat such as hidden structure defects and detection of underground utilities especially pipes or cables. It can act as a safety aide since it is used to avoid potential and dangerous problems for the detection such as buried explosive object and treatment of hidden problems, for example, concrete structure defect. From the previous research there are many analyses of GPR data using features extraction and machine learning algorithms. All this research is done on different objects such as underground pipes [1][2], landmines [3], urban roads [4], crevasses [5], or concrete structure [6]. Aleksandar Bugarinović and his team [7], the extraction of features from hyperbolic patterns is conducted through the discrimination of segments of interests (SOI). A basic pre-processing of radargram is done and the detection of SOI in dense radargram is done using the Cascade Object Detector. It is trained based on sets of samples that relevant to the presence of hyperbolic reflections. Extraction of data such as the points of apices, points of prongs are determined, points of intersections between hyperbola patterns are also checked. Two algorithms had been used as comparison which is Baseline CNN and traditional learning-based algorithm. In this experiment, the Baseline CNN successfully shows that the localization of coordinates of apices and prongs, and

recognition of the interferences of two or more hyperbolic reflections. Therefore, the CNN model is shown to outperform the traditional learning algorithm. Umut Ozkaya et al. suggest a Convolutional Support Vector Machine (CSVM). CSVM is like CNN in that it has a cascade of convolution and pooling layers. The primary distinction is that it produces feature maps using linear Support Vector Machine (SVM)-based filters and employs a forward learning method. Umut Ozkaya and colleagues utilised the proposed method to categorise buried things, form kinds, and soil types.[8] A Discrete Wavelet Transform-Fractional Fourier Transform (DWT-FRFT) featured-based SVM system is used to classify the material of the underground objects [9]. The features extraction method is compared with statistical features and frequency domain features, with the same SVM as classifiers for evaluation of the performance. The proposed method outperformed the other methods and with results of more than 90.0%. GPR data is time-consuming, and yet is heavily relying on the labour-intensive work of well-trained experts [10]. Furthermore, the GPR data would be extremely affected by background noise [11]. Therefore, the manpower will face difficulties in classifying targeted objects that does not have huge differences especially from shapes and noisy environment which can be easily ignored [12]. As the consequences, the amount of information and requirement of complex interpretation are becoming the main reasons for developing the automated steps and process for analysis of anomalies in radargram [13]. Li Liu et. al. propose a novel underground object classification algorithm using deep 3-D convolutional networks (C3D) and multiple mirrors encoding (MME) for 3-D GPR data. They apply 3-D GPR data as training data for C3D to capture the spatio-temporal features between parallel B-scans to improve the classification accuracy. The results showed that the proposed method outperforms the state-of-the-art B-scan-based methods [14]. Mostafa Elsaadouny et. al. investigates the use of a small dataset to implement the LeNet-5 convolutional network (ConvNet) for image classification. The network performance was monitored, and the classification results demonstrated a high level of precision and accuracy [15]. Ibrahim Mesecan uses synthetic data from GprMax simulation software and proposes a 3-step method to locate and discriminate underground objects. The proposed method is tested using 2 classifiers. Result shows the overall performance increases comparably for Support Vector Machines (SVM) from 67.6% to 85.5% [16].
Numerous research has been conducted in this area, but more work must be done in terms of predicting and classifying the size of buried object. N. R. Syambas largely focuses on the size of fundamental object with three different shape objects. Additionally, the approach only employs one classifier and not compare to another classifier [17]. Nairit Barkataki et. al. focused on prediction size of cylindrical buried object and use two different classifiers to compare the performance, but the data samples are created using GPR simulator device [18]. This article is proposing a framework to classify the size of the underground objects focussing the object at different depth using real GPR device namely MALÅ CX Concrete Scanner. This paper also focuses on classification of size of buried object from GPR data using both Backpropagation Neural Network and Support Vector Machine classifier. The test result of both classifiers is then compared to evaluate the proposed approach's performances.

2. Methodology

The aim of this project is to classify the sizes of the underground metallic pipe according to the hyperbolic reflections from the GPR data. These categories are separated to two size categories namely the diameters of 3.5cm and 4.8 cm respectively. The results obtained should contain the accuracy of classification based on respective machine learning algorithms. The process above is summarize as flow chart in Figure 1.

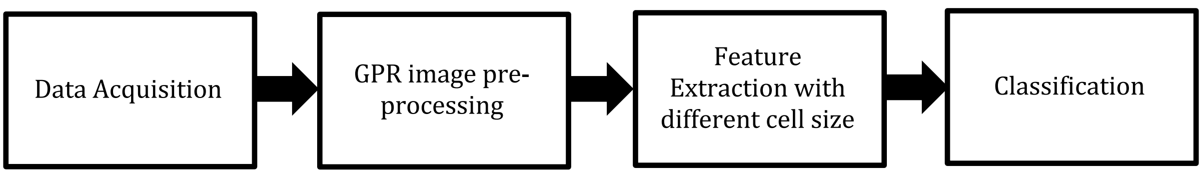

Figure 1. Flow chart of the proposed process.

2.1. Data Acquisition

The experimental setup is prepared as shown in Figure 2. The underground metallic pipes with different diameters and buried with respective depth. These samples were then buried in the sandbox test bed with two different depths namely at 10 cm and 15 cm from the top surface individually. The soil medium is dry sand as it has different permittivity compared to the metallic pipes. The sand also been used because it is one of the best medium that can shape the hyperbolic signature ideally. Underground pipes are metallic and have diameters of 3.5 cm and 4.8 cm respectively. Pipes with a diameter of 3.5 cm are classified as Group 1, those with a diameter of 4.8 cm as Group 2. The GPR direction of scanning is illustrated in Figure 3.

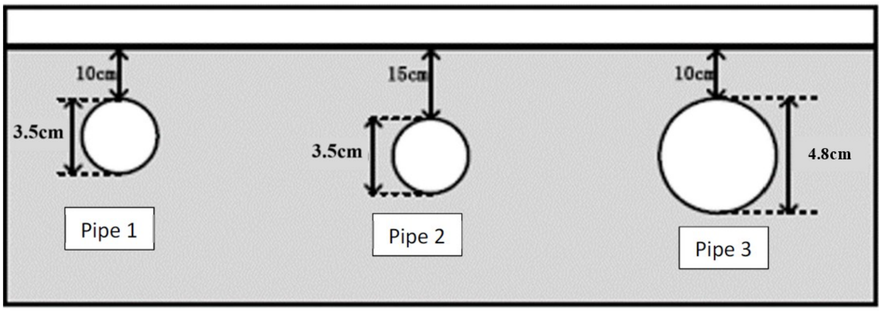

Figure 2. The experiment setup to be scanned using GPR.

The experiment was constructed at a sand medium of test bed (fully wooden crates without any metal like nail for assemble) as shown in Figure 4. It is because to avoid from the signal interference when collecting data occur. MALÅ CX Concrete Scanner is used to collect the data of hyperbolic pattern the buried object in this project. It consists of a few components such as 1.2GHz antenna, Li-Ion battery and encoder. In this project this equipment use to get the b-scan of buried object by electromagnetic wave. The mapping and scanning operation of GPR system were supervised by the authorized officer of Non-Destructive Test-Material Structure Integrity Group at Malaysia Nuclear Agency. The coding "readmala" in MATLAB is used to interpret and analyse the raw data to acquire the GPR image in B-scan format as shown in Figure5. It demonstrates the wave's reflection and is returned to the GPR instrument. A total of 233 raw data is recorded based on the experimental setup mentioned. In case there is any raw data that does not show output that is like Figure 5, the raw data file will be excluded to prevent any misclassified data from mixing into the data set.

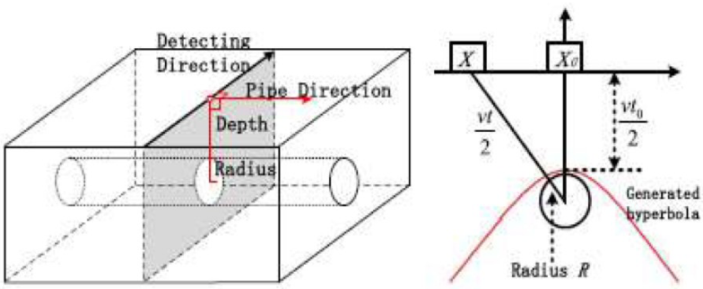

Figure 3. The direction of GPR scanning.

Figure 4. Sandbox test bed.

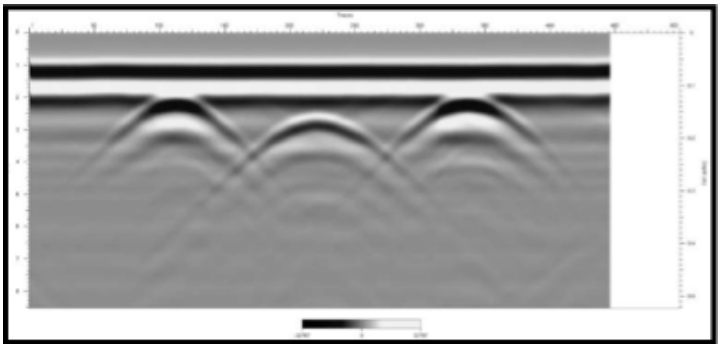

Figure 5. The example of GPR data.

2.2. Data Pre-processing
Before the raw data is pre-processed, the raw data is sorted out to exclude any unimportant data. An amount of 204 images is kept and the pre-processed stage is continued. Firstly, the image read is converted to grayscale image using MATLAB software. Figure 5 depicts a grayscale image in which the colour black represents the object (wave and surface), and the colour white represents the surrounding area.
After the conversion, the contrast of images is enhanced to show region and boundaries of the hyperbolic curves. In addition, the next process is completed by resizing the image by resize factor of 2 that indicates that the image has been enlarged. Therefore, the process of cropping the images can be made easier. The process continues with finding the specific coordinates of the desired region. Cropping the GPR image allows the image to focus on the GPR's hyperbolic signature. It can also reduce noise and unwanted objects in the GPR image's hyperbolic signature. Figure 6 shows an example of a cropped image. After deciding the region and cropped out the image, the image will be rescaled again to standardize the image size. This is due to the fact that all of the data must be combined in the same resolution in order for the feature extraction and classification processes to be error-free. The final output of the image pixel is 150 in height and 230 in width.

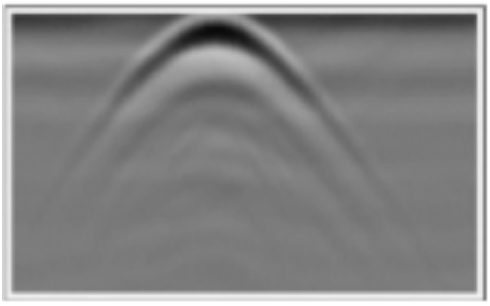

Figure 6. Cropped image

2.3. Feature Extraction
In this research, image processing method for underground object size classification is used by extracting the feature from hyperbolic image from GPR data using the histogram of oriented gradient (HOG). This method divides an image into its constituent local parts and generates identifiers for each one. An image is a series of local regions, according to the HOG algorithm. Each local histogram represents a distribution of the number of measured gradients that occur in each direction in a particular local area known as "cell." The feature extraction procedure in the HOG algorithm consists of three primary stages namely:
1. calculation of the magnitudes and directions of gradients.
2. gradient histograms are created for each cell.
3. normalising block-by-block cell histograms with each other.

The magnitudes of the horizontal and vertical gradients, m, and their directions, θ, are computed using equations (1) and (2) where g_x and g_y are the vertical and horizontal components of gradient respectively.

$$m(x,y) = \sqrt{g_x(x,y)^2 + g_y(x,y)^2} \quad (1)$$

$$\theta(x,y) = tan^{-1}\frac{g_y(x,y)}{g_x(x,y)} \quad (2)$$

During HOG features extraction, three different cell size of [16*16], [30*30] and [40*40] are used for extracting different number of features vectors. The number of feature vectors extracted are 3744, 540 and 288 for respective cell size. The feature vectors will undergo normalization to change the values of dataset in a common scale without distorting the differences in the range of values. Figure shows the HOG image after undergo feature extraction process using cell size of [16*16].

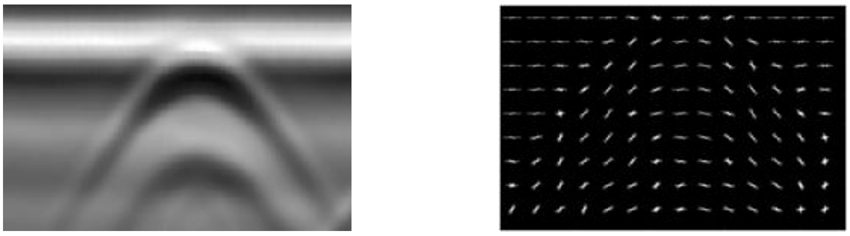

Figure 7. Example of conversion from pre-processed GPR image to HOG image.

2.4. Classification
SVM is a supervised model with a learning algorithm linked with it that is primarily used for data analysis for regression and classification. The SVM searches for the optimum decision border between the boundaries of the two classes. The margin is maximum for and is closest to the decision boundary for the boundary that is chosen as the decision boundary. SVM is used to categorise the size of underground object. Prior to testing, SVM must be trained. The data set proportion is set in three different percentage for training to testing set which is 60:40, 70:30 and 80:20. To validate the results, cross-validation method is also carried out during training progress.

To compare and evaluate the results obtained, Backpropagation Neural Network is also used. for Backpropagation Neural Network, the number of hidden layers must be determined, and the number of hidden layers decided is 15 with 3744, 540 and 288 features as input. The process of classification is repeated three times and the average value of respective aspects such as computation time, training accuracy and testing accuracy is obtained and tabulated as result

3. Result
In this study, accuracy to classify size of underground object is tested using SVM and Backpropagation neural networks. From this project, the success of classifying the size of underground pipe is shown for all classifiers. HOG is used together with SVM as the feature vectors extracted from HOG method is perfectly suits for SVM in classification and regression problems especially in non-linear cases. In this work, there are three sets of data set which is 3744, 540 and 288 features. 408 rows of sample are prepared, and all the classes are labelled in last column respectively.

From Table 1, except for the data set with 288 features, all the other data set obtain 100.00% training accuracy. Although the training accuracy does not reach perfection, the range of average training accuracy still in range of 98.78% to 99.65% which is only has minor misclassification issue. The data set with 288 features might not consists of enough features for the SVM to undergo classification accurately. For the data set of 540 features, all data set proportion has obtained 100.00% for average testing accuracy. Although the data set of 3744 features has 100.00% training accuracy, it may contain insufficient information as well for certain data which leads to the imperfection in the 40% testing model. For the data set of 288 features, the testing accuracy is 100.00% in the training to testing set of 80:20 which has less amount of testing data. The computation time for data set of 540 and 288 features has nearer range from 1.33 second to 1.49 seconds compared to that of 3744. Therefore, the data set mentioned is preferable compared to data set with 3744 features. However, the computation time in SVM is expected to be shorter as it is only needed to classify the results only without extracting features during the process.

Table 1. Accuracy result from SVM.

Number of features	Data set Proportion (%)		Average Computation Time, t (s)	Average Training Accuracy, Tr (%)	Average Testing Accuracy, Te (%)
	Train	Test			
3744	60	40	2.71	100.00%	99.40%
	70	30	3.18	100.00%	100.00%
	80	20	3.40	100.00%	100.00%
540	60	40	1.41	100.00%	100.00%
	70	30	1.46	100.00%	100.00%
	80	20	1.43	100.00%	100.00%
288	60	40	1.33	98.78%	98.18%
	70	30	1.49	99.65%	99.18%
	80	20	1.36	99.40%	100.00%

Table 2. Accuracy result from Backpropagation Neural Network.

Number of features	Data set Proportion (%)		Average Computation Time, t (s)	Average Training Accuracy, Tr (%)	Average Testing Accuracy, Te (%)
	Train	Test			
3744	60	40	3.93	100.00%	100.00%
	70	30	3.56	100.00%	100.00%
	80	20	3.48	100.00%	99.20%
540	60	40	2.87	100.00%	100.00%
	70	30	2.99	100.00%	100.00%
	80	20	3.02	100.00%	100.00%
288	60	40	2.98	100.00%	100.00%
	70	30	2.80	100.00%	100.00%
	80	20	2.77	100.00%	100.00%

In Table 2, the computation time shown for backpropagation Neural Network is in range of 2.77 second to 3.93 second. In overall comparison, the time taken for classification is slightly longer compared to SVM. However, the duration of the time is still increasing according to the number of features in data set. Besides that, the duration is affected by the training mode of the neural network which is scaled conjugated gradient that requires less memory and therefore the computation time will be reduce. All the data set has achieved training accuracy of 100.00% which is considered as great performance. It indicates that the features extracted can be completely classified in this neural network. From Table 2 it is shown that, except for the data set of 3744 features with training to testing set of 80:20, other trials had achieved 100.00% testing accuracy.

The experimental results proof that this classifier was successful in classifying the pipe size because the accuracy results are all above 98%. This demonstrates that the classifier can produce accurate predictions using the features it has gathered. Overall, this trial result demonstrates the system is efficient to classifying the underground object size.

4. Conclusion

This project is mainly classifying the size of underground metallic pipe with same material into respective category which is Group 1 for diameter of 3.5 cm and Group 2 for diameter of 4.8 cm. A total amount of 204 images and samples are used to classify and all these data set has gone through pre- processing stages, HOG feature extractions and classification which consists of SVM and Backpropagation Neural Network. The classification for SVM and Backpropagation Neural Network obtaining great classification results. The data set used is feature vectors from HOG feature extractions algorithm. As a conclusion, objectives of this project are achieved. In SVM and Backpropagation Neural Network, both classifiers had shown the great performance by using the feature vectors.

However, using a smaller number of features, classification accuracy using Backpropagation Neural Network shows better performance compare to SVM. The optimum number of features will be great for classification results. All these variables required in HOG features extraction algorithm still requires knowledge in this algorithm to decide the optimum cell size. Future effort the classes of the data can be added with other categories of underground metallic pipes. Besides using 3.5 cm and 4.8 cm diameters, other measurement of the diameter such as 6.0 cm or any comparable sizes can be added. Besides that, the setup of environment can be changed by added with different objects such as rocks to produce GPR data with noise.

5. References

[1] M. Ristić, Aleksandar Bugarinović, Željko Vrtunski, Milan Govedarica, "Point coordinates extraction from localized hyperbolic reflections in GPR data," J. Appl. Geophys., vol. 144, pp. 1–17, Sep. 2017, doi: 10.1016/j.jappgeo.2017.06.003.

[2] H. Bai and J. V. Sinfield, "Improved background and clutter reduction for pipe detection under pavement using Ground Penetrating Radar (GPR)," J. Appl. Geophys., vol. 172, p. 103918, Jan. 2020, doi: 10.1016/j.jappgeo.2019.103918.

[3] M. Moalla, H. Frigui, A. Karem, and A. Bouzid, "Application of Convolutional and Recurrent Neural Networks for Buried Threat Detection Using Ground Penetrating Radar Data," IEEE Trans. Geosci. Remote Sens., vol. 58, no. 10, pp. 7022–7034, Oct. 2020, doi: 10.1109/TGRS.2020.2978763.

[4] B. Park, J. Kim, J. Lee, M.-S. Kang, and Y.-K. An, "Underground Object Classification for Urban Roads Using Instantaneous Phase Analysis of Ground- Penetrating Radar (GPR) Data," Remote Sens., vol. 10, no. 9, p. 1417, Sep. 2018, doi: 10.3390/rs10091417.

[5] B. Walker and L. Ray, "Multi-Class Crevasse Detection Using Ground Penetrating Radar and Feature-Based Machine Learning," in IGARSS 2019 - 2019 IEEE International Geoscience and Remote Sensing Symposium, Jul. 2019, pp. 3578– 3581, doi: 10.1109/IGARSS.2019.8899148.

[6] L. Jiao, Q. Ye, X. Cao, D. Huston, and T. Xia, "Identifying concrete structure defects in GPR image," Measurement, vol. 160, p. 107839, Aug. 2020, doi: 10.1016/j.measurement.2020.107839.

[7] M. Ristić, Aleksandar Bugarinović, Željko Vrtunski, Milan Govedarica, "Point coordinates extraction from localized hyperbolic reflections in GPR data," J. Appl. Geophys., vol. 144, pp. 1–17, Sep. 2017, doi: 10.1016/j.jappgeo.2017.06.003.

[8] U. Ozkaya, F. Melgani, M. Belete Bejiga, L. Seyfi, and M. Donelli, "GPR B scan image analysis with deep learning methods," Measurement, vol. 165, p. 107770, Dec. 2020, doi: 10.1016/J.MEASUREMENT.2020.107770.

[9] Q. Lu, J. Pu, and Z. Liu, "Feature Extraction and Automatic Material Classification of Underground Objects from Ground Penetrating Radar Data," J. Electr. Comput. Eng., vol. 2014, pp. 1–10, 2014, doi: 10.1155/2014/347307.

[10] B. Park, J. Kim, J. Lee, M.-S. Kang, and Y.-K. An, "Underground Object Classification for Urban Roads Using Instantaneous Phase Analysis of Ground- Penetrating Radar (GPR) Data," Remote Sens., vol. 10, no. 9, p. 1417, Sep. 2018, doi: 10.3390/rs10091417.

[11] J. Feng, L. Yang, H. Wang, Y. Song, and J. Xiao, "GPR-based subsurface object detection and reconstruction using random motion and DepthNet," arXiv. pp. 7035–7041, 2020.

[12] Q. Lu, J. Pu, and Z. Liu, "Feature Extraction and Automatic Material Classification of Underground Objects from Ground Penetrating Radar Data," J. Electr. Comput. Eng., vol. 2014, pp. 1–10, 2014, doi: 10.1155/2014/347307.

[13] J. Baili, S. Lahouar, M. Hergli, I. L. Al-Qadi, and K. Besbes, "GPR signal de- noising by discrete wavelet transform," NDT E Int., vol. 42, no. 8, pp. 696–703, Dec. 2009, doi: 10.1016/j.ndteint.2009.06.003.

[14] Liu, L., Yu, H., Xu, H., Wang, B., & Li, J. (2022). Underground Object Classification Using Deep 3-D Convolutional Networks and Multiple Mirror Encoding for GPR Data. *IEEE Geoscience and Remote Sensing Letters*, *19*, 1-5.

[15] Elsaadouny, M., Barowski, J., & Rolfes, I. (2020, March). Extracting the features of the shallowly buried objects using LeNet convolutional network. In *2020 14th European conference on antennas and propagation (EuCAP)* (pp. 1-4). IEEE.

[16] Meşecan, İ., Çiço, B., & Bucak, İ. Ö. (2020). Feature vector for underground object detection using B-scan images from GprMax. *Microprocessors and Microsystems*, *76*, 103116.

[17] Nana Rachmana Syambas, "An Approach for Predicting the Shape and Size of a Buried Basic Object on Surface Ground Penetrating Radar System", International Journal of Antennas and Propagation, vol. 2012, Article ID 919741, 13 pages, 2012. https://doi.org/10.1155/2012/919741

[18] Barkataki, Nairit & Mazumdar, Sharmistha & Talukdar, Rajdeep & Chakraborty, Priyanka & Tiru, Banty & Sarma, Utpal. (2020). Prediction of Size of Buried Objects using Ground Penetrating Radar and Machine Learning Techniques. 781-785. 10.1109/ComPE49325.2020.9200094.

Acknowledgments
This project has been supported by RESMATE 2020 UniMAP Internal Grant. This project also supported by Nuclear Agency Malaysia by providing GPR RAMAC/MALA system under MoU agreement.