L'ÉNERGIE AU CŒUR DES ONDES
RESSOURCES ET ENVIRONNEMENT :
GESTION "INTELLIGENTE"

# Multi-label classification with deep learning techniques applied to the Bscan image on radar GPR

*El Karakhi Abouzouhour Soukayna, Reineix Alain, and Guiffaut Christophe*
*University of Limoges, XLIM Laboratory, Limoges, France*

*Keywords: Ground Penetrating Radar,Image processing, Buried objects,Deep learning;*

## Abstract

The ground penetrating radars (GPR) are now widely used for the detection of buried objects in areas such as: geology, archaeology and civil engineering. It has the advantage of allowing detection by a non-destructive technique. The principle for time domain GPR consists in emitting electromagnetic pulses in the ground, these one are then reflected by the targets to be detected. A single GPR signal trace captured at a position of the radar is a 1D signal called Ascan. A set of Ascan radar waveforms captured at a certain number of different consecutive positions along a particular direction will form a 2D signal called B-scan in the case of a rectilinear displacement. They show response shapes of hyperbolic type and its analysis gives many characteristics. For example, in the case of buried pipes, a specific processing allows to find the diameter of these pipes, their nature as well as the characteristics of the ground. However, these approaches often require complex post-processing of the Bscan, which can be time-consuming and therefore makes it difficult to perform real-time characterisation at the expense of such methods. With the emergence of deep neural networks and with a learning phase on a large number of Bscan, it becomes possible to extract almost instantaneously the characteristics of GPR radar data. In this study, a multi-label classification (MLC) model based on transfer learning and data augmentation was developed to generate multiple information elements on the same image and to realize classification. Three deep learning models: VGG-16, ResNet-50 and adapted CNN were used as pre-trained models for transfer learning. The networks were trained on the dataset created in this study and evaluated on a set of performance metrics.

## Résumé

Les radars à pénétration de sol (GPR) sont maintenant largement utilisés pour la détection d'objets enterrés dans des domaines tels que la géologie, l'archéologie et le génie civil. Il présente l'avantage de permettre une détection par une technique non-destructive. Le principe du GPR temporel consiste à émettre des impulsions dans le sol, celles-ci sont ensuite réfléchies par les cibles à détecter. La formation d'une image à partir des différentes traces (Ascan) mesurées pour différentes positions du radar est appelée Bscan dans le cas d'un déplacement rectiligne. Ils présentent des formes de réponse de type hyperbolique et leur analyse donne de nombreuses caractéristiques. Par exemple, dans le cas de canalisations enterrées, un traitement spécifique permet de retrouver le diamètre de ces canalisations, leur nature ainsi que les caractéristiques du sol. Cependant, ces approches nécessitent souvent un post-traitement complexe du Bscan, qui peut être long et rend donc difficile la réalisation d'une caractérisation en temps réel. Avec l'émergence des réseaux de neurones profonds et avec une phase d'apprentissage sur un grand nombre de Bscan, il devient possible d'extraire quasi instantanément les caractéristiques des données radar GPR. Dans cette étude, un modèle de classification multi-label (MLC) basé sur l'apprentissage par transfert et l'augmentation des données a été développé pour générer plusieurs éléments d'information sur une même image et réaliser une classification. Trois modèles d'apprentissage profond : VGG-16, ResNet-50 et CNN adapté ont été utilisés comme modèles pré-entraînés pour l'apprentissage par transfert. Les réseaux ont été entraînés sur le jeu de données créé dans cette étude et évalués sur un ensemble de mesures de performance.

## 1 Introduction

Ground Penetrating Radar (GPR) is an increasingly significant and effective approach to conducting non-destructive engineering investigations. Over an extended time span, the GPR has been applied extensively to geological research, in civil engineering, as well as in agriculture and environment. Recently, different approaches have been proposed by scientists for the processing of GPR data. T. Noreen [1] suggests a machine learning based approach for hyperbolic signature identification using a SVM (support vector machine) with the histogram of oriented gradient features (HOG). E. Temlioğlu [2] proposes different approaches in landmine detection: among them Binary Robust Independent Elementary Features (BRIEF), Edge Histogram Descriptor (EHD), Histogram of Oriented Gradients (HOG), Scale Invariant Feature Transform (SIFT) and Speeded Up Robust Features (SURF). In this study, W. A. Wahab [3] suggested a novel hyperbola fitting technique for estimating the radius of buried utilities (pipes and cables). The approach was implemented on buried pipes of nine different diameter values. B. Walker and L. Ray have implemented a feature-based machine learning

approach to processes GPR data. The researchers employed feature vectors from the Histogram of Oriented Gradients (HOG) with a Support Vector Machine (SVM) to detect deep and shallow crevices. Additional characteristic operators are also used as a pre-processing approach, such as Sobel's,[5] Wavelet Edge Detection, [6] Canny's operators. These approaches are very slow in computation, but they provide very accurate and precise extraction of the characteristics of the B-Scans.
In recent years, several researchers have proposed approaches for the automatic identification and localization of buried objects based on deep learning models. This has demonstrated the many benefits of deep learning technologies. As the amount of GPR data expands, the capabilities of the existing machine learning techniques are becoming less performant for digital image processing. Researchers [7] [8] [9] [10] increasingly began to employ CNN-based methods of learning hyperbolic shapes for the classifications and identification of buried objects. The majority of the methods mentioned here focus on identifying and positioning buried objects, and classifying them according to a single characteristic. Recently, very limited studies have focused on a multi-label classification model to provide multiple pieces of information from a single image.
The paper's contributions are synthesized as described below:

- Database creation based on the Finite Difference Time domain (FDTD) simulation,
- Image processing using gradient operators for edge region detection ( Sobel, Canny, and Prewitt),
- Development of a multi-label model for the identification of the diameter of buried pipes, their inner space filling as well as the identification of the soil medium based on three models: VGG 16 - Resnet 50 - and customized CNN.

The originality of this work concerns the development of deep learning model architectures, and the approach to model prediction and analysis of the three models developed.

## 2 Methodology

In order to generate our dataset, we will use the TEMSI-FD software developed within the EMC team of XLIM Institute. It consists of a software FDTD-based method [12] designed to simulate the propagation of electromagnetic waves in complex media. For the GPR study, the time domain signal emitted, propagating and received by antennas can be calculated. It is useful to represent these recordings in two forms: the A-scan (1D response), the B-scan (2D image resulting of multiple consecutive A-scan) on the figure 1.

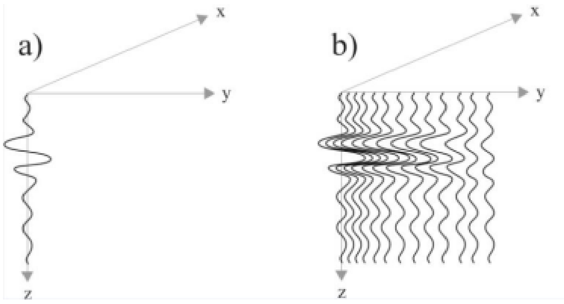


Figure 1: (a) A-scan GPR signal (b) B-scan GPR signal

To create a simulation model, a number of settings are required. First, the entire scene has to be designed. The scene is characterized by the propagation medium of the electromagnetic waves. The model of the ground is based on the fractal model [11], but also on homogeneous media such as dry clay, dry sand and concrete. The dimensions of the scene in cells number is 251 x 200 x301 (X x Y x Z), the cell size being 4mm x 4mm x 4mm. The soil layer depth is 1m. It is the maximum depth value to be probed by the GPR, and consequently, it determines the estimation of the maximum time of observation/simulation, which is 20 ns for this volume. This duration was calculated for the case of the concrete layer which has the smallest value for the propagation velocity among the media considered: $v_p = c/\sqrt{\varepsilon_r}$ with c the velocity of light. The GPR antennas system is composed of a fixed emitter and 10 receiving antennas. The receivers are also spaced 20 mm apart and are moved in groups of 200 mm along the horizontal scan line on the soil, reducing the number of simulations by 10. The multi-receiver bistatic GPR is illustrated in the Figure 2:

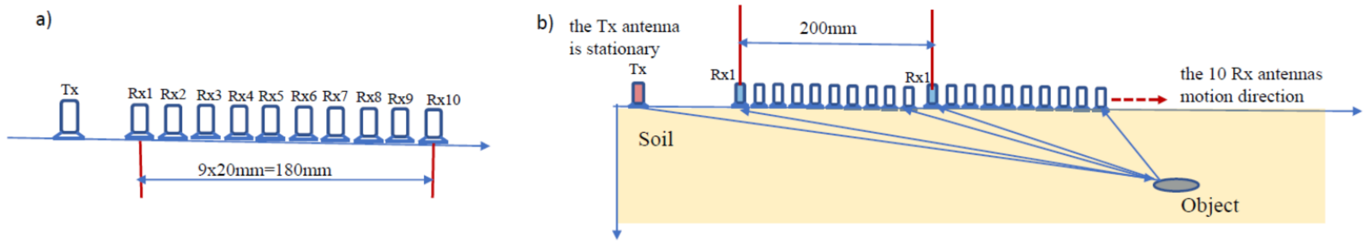


Figure 2: a) The multi-receiver 10 Rx antennas bistatic GPR with one Tx antenna emitter architecture, b) The way the 10 Rx antennas are moved in one direction during the exploration on the soil

A set of 5400 GPR B-scans database concerning 4 soil types, 50 depth values and 9 different pipe diameters is generated; the current diameters for the simulation data are 16, 24, 32, 40, 48, 64, 72, 80 and 100 mm. The depth values range from 204 mm to 400 mm below the soil surface. Three types of pipes were chosen during the simulation: a metal pipe, an air-filled pipe, and a water-filled pipe.

The table 1 shows the dielectric parameters for the soil types:

| SI. No | Soil type | Conductivity(S/m) | Relative permittivity $\varepsilon_r$ |
|---|---|---|---|
| 1 | Dry sand | 0.002 | 10.00 |
| 2 | Dry clay | 0.001 | 5.53 |
| 3 | Concrete | 0.001 | 8 |
| 4 | Fractal | 0.001 | 6 |

Table 1: Dielectric parameters for the soil type used in the simulation

The relevant parameter values used in the simulation are indicated in Table:

| Parameters of simulations | Values |
|---|---|
| Source frequency | 1.00 Ghz |
| Source waveform | sino-gaussian pulse |
| Source pulse duration | 1ns |
| A-scans intervals | 20mm |
| No of A-scans | 40 |
| Spatial resolution | 4mm |

Table 2: The parameters values of simulations

To improve the quality of the measurements, it is necessary to eliminate the surface echo and the direct coupling between the antennas and thus keep only the useful signal. In our case, we proceed to a second recording with a scene that contains no objects and then we subtract this second measurement from the one obtained with a scene containing the objects. The figure 3 is illustrating the resulting treatment.

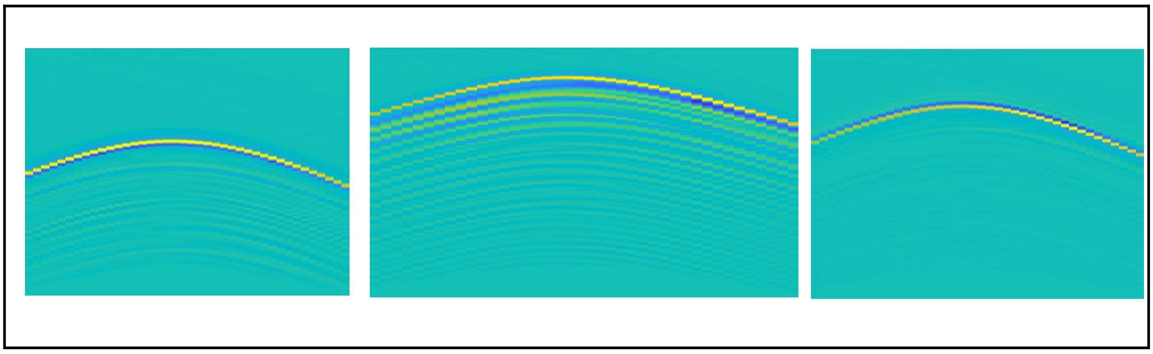

Figure 3: (a) Diameter 64mm, metal, concrete, (b) Diameter 16mm, water filled pipe, dry clay, (c) diameter 80 mm, empty pipe, dry sand

Now we will describe the CNN and the procedure used to apply deep learning for retrieving the soil and pipes characteristics. First, the global GPR data handling workflow in this study involves feature extraction, splitting the dataset into training and test sets, CNN model training and prediction. The feature extraction pipeline for B-Scan images consists of testing several approaches to edge detection. The goal of this technique is to significantly reduce the amount of data while retaining information that can be considered more relevant. The objective is to focus on regions of the hyperbola where there is a distinct intensity variation. The hyperbola region is detected by a bounding box, the image is cropped to extract the region of interest from each frame. For this purpose, we have focused on Canny's operators. Canny is a multi-step edge detection algorithm developed by John F. Canny in 2001 [13], which proceeds through the following steps:

- Noise reduction: Noise is removed from the image with a 5x5 Gaussian filter.
- Determine the gradient intensity of the image using the Canny operator.
- Non-maximum Suppression: This step consists of eliminating those undesirable pixels that may not be edges.
- Hysteresis Thresholding

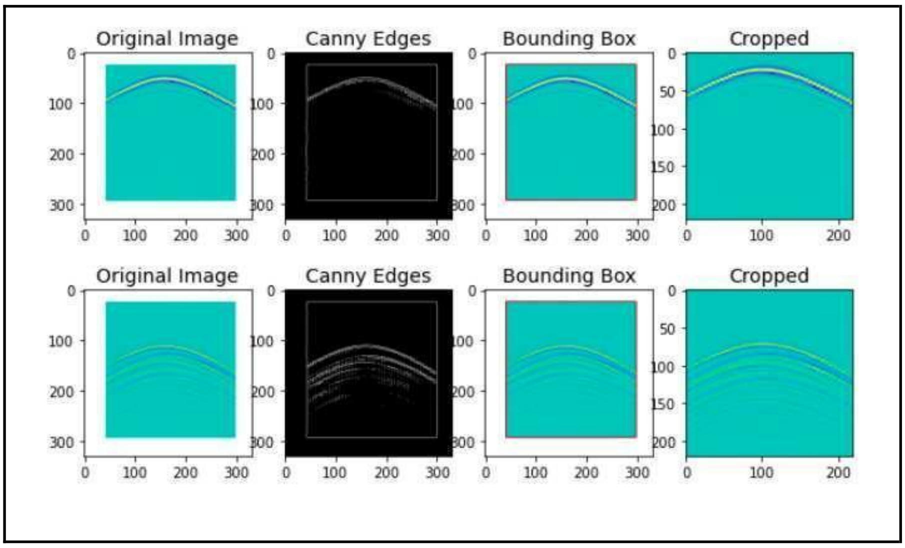


Figure 4 : Example of output from edge detection operators

## 3 Proposed Deep Learning architecture:

The objective of this work is to implement Deep Learning models to extract features like buried pipes diameters or type of soil layer from B-scan images. We will implement three different architectures: VGG-16, Resnet-50 through learning transfer and a custom CNN model. The proposed architecture is a multi-label model classification in which individual objects can be classified into multiple classes at one time, compared to traditional one-label classification cases involving a single class of objects. Multi-label classification approaches are becoming unavoidable in advanced technology. In this work, in the classification of B-Scan images, each image can belong to several distinct categories. For example, we classify the type of pipe (metal pipe, water-filled pipe and empty pipe), 9 different diameters, so that we distinguish four different mediums. As seen above, the dataset contains 5400 B-Scans, it will be divided in such a way that 80% is reserved for model training, and 20% for test data. Now we will describe the different topologies.

### 3.1 Transfer learning using VGG 16 and Resnet 50

Residual neural networks: ResNet, an acronym for Residual Networks, is a classical neural network used as a backbone in many computer vision tasks. The major revolution of ResNet is its ability to successfully train extremely deep neural networks with more than 150 layers. Before ResNet, training very deep neural networks was difficult due to the gradient backpropagation problem. The strength of this architecture lies in the concept of skip connection that is illustrated in the figure above. On the left of the figure 5 we stack the convolution layers successively. On the right, we stack the convolution layers as before, but we connect the original input to the output of the convolution block

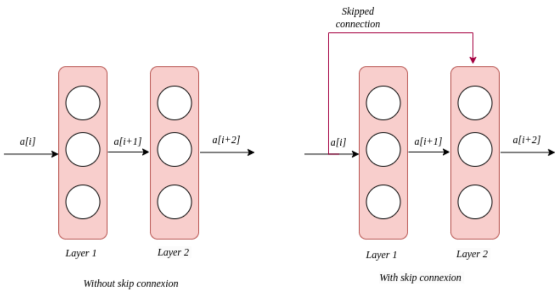


Figure 5 : Residual learning: a building block.

Visual Geometry Group (VGG-16): VGG stands for Visual Geometry Group; it is a standard deep Convolutional Neural Network (CNN) architecture with multiple layers. VGG introduced a sample and effective design by using a stack of 3x3 filters and demonstrated that placement of small kernel size (3x3) filters could induce the effect of the large filter size while providing more room for increasing the depth of the neural network. VGG models emphasized the importance of depth in convolutional neural networks, as a deep network with small filter sizes can learn a larger number of filters and a more complex representation of the data. This approach was a significant departure from earlier models that relied on large filters to capture low-level features, and VGG models were able to demonstrate that small filters could effectively capture both high-level and low-level features. VGG-16 uses 16 trainable layers in their network which are shown in Figure 6.

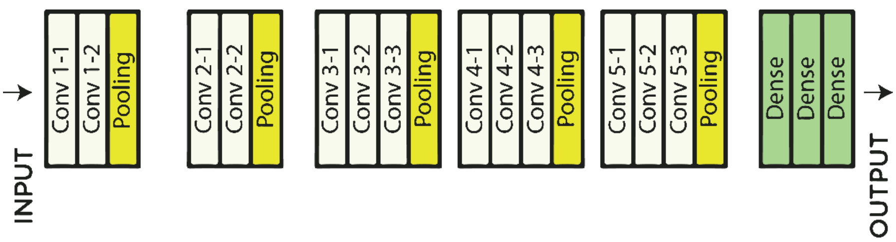


Figure 6: VGG-16 architecture

In this study, we use these as initial weights in our model and start training on our dataset with these weights. The last layer has been replaced by three ones: the first branch represents the diameter classification layer, the second one represents the pipe type classification and the last one the different of propagation of medium. Figure 7 summarizes the concept of transfer learning.

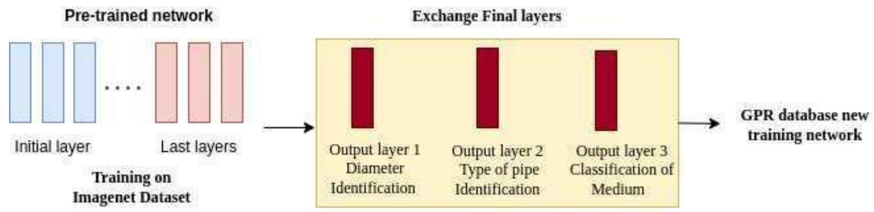


Figure 7: Transfer learning strategy

### 3.2 Custom Model

In this work, a deep convolutional neural network is developed for multi-label classification to classify 9 different diameters, 3 different pipe types, and 4 different propagation media. The custom CNN model is inspired by the inception network, it is a model developed by google, the main difference between the inception model and the ordinary CNNs are the inception blocks. These consist in convoluting the same input with several filters and concatenating their results. The aim of this model is to introduce the concept of multipathing, which allows different characteristics to be simultaneously captured and extracted. The building block is illustrated in the figure 8, this block contains different convolution layers of different filter sizes (1x1), (5x5), (7x7). The diversity of the kernel size will increase the capacity of the network to extract the most complex features. 32 filters of size 3 × 3 are used in the input layer, followed by a MaxPool layer and then a batch normalization layer to increase the speed and stability of learning, followed by a convolution layer with 64 filters of size 3x3, after which a max pooling layer was defined, followed by two inception building blocks represented above, each convolution layer contains 128 and 16 filters respectively of size (1x1), (5x5), (7x7), at the end of each inception block, all feature maps are concatenated. The activation function employed in all convolutional layers is the rectified linear unit (ReLU) defined as: The output of ReLU is equal to 0 if the input value is less than 0 and equal to the input value if the input is positive. Then, all the feature maps are transformed into a 1D vector using a flattened layer. A set of 3 dense layers has been defined, containing 256, 512, and 256 neurons respectively. Each layer is followed by an activation function Relu. Finally, 3 branches were defined for each output, each branch includes upstream a set of three dense layers respectively consisting of 96, 48 and 48 neurons, each output layer has 9 neurons to recognize the diameter size, 4 neurons to specify the propagation medium, and 3 neurons to distinguish the pipe type. The output layers are followed by a softmax layer mainly used for multi-label classification problems to predict the probability of each label.

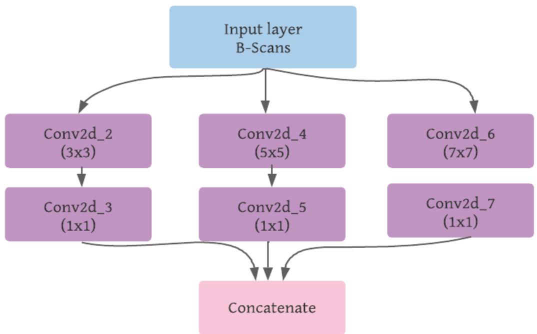


Figure 8: Custom CNN Model - building block

## 4 Results and discussion:

Three indicators were used to evaluate the three Deep Learning models: Precision, Recall, and F1- score, the accuracy is the simplest indicator, it measures the percentage of correct predictions. It is defined as follows:

$$accuracy = \frac{TP+TN}{TN+TP+FP+FN} \tag{1}$$

Based upon the results presented in Figure 9, we observe a better accuracy of 96% for the diameter identification for the VGG16 model, and an efficiency of 98% for the pipe type identification achieved by the Resnet50 model.

We also reach a higher level of sensitivity for the identification of the propagation medium, which achieves 98% for both the VGG16 model and the customized CNN model.

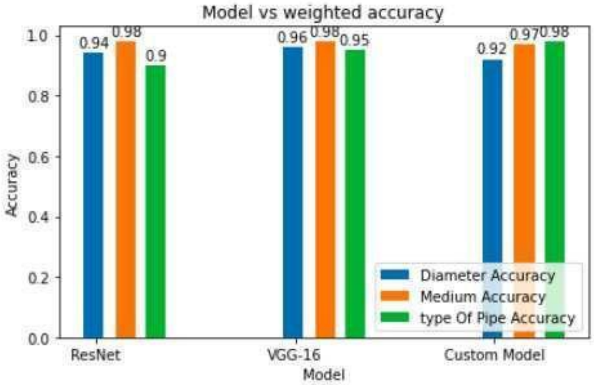


Figure 9: Accuracy vs. weighted accuracy of the three developed models with Canny's operators

TP (True positive) stands for the number of accurate diameter, pipe type and medium identifications, FP (False Positives) represents the number of incorrect identifications and FN (False negative) represents the number of failed recognitions. In GPR, accurate identifications mean that the target is well located and accurately classified. Incorrect identifications are indicative of a well-located target, but its geometry, as well as its type and medium, are not properly classified. The missed identifications reflect the target not being located. The table 3 shows the performance of the model Resnet-50 based on the 9 classes for the identification of diameters. The overall accuracy is 96%. Table 3 shows the classification report for the 9 diameter values, we find the precision that indicates the performance or the positive prediction given by the model, recall is a statistical metric that shows how many positive cases actually correspond to the predicted class. The F1 score provides the information about the incorrect predictions of the model meaning that 1 is the best and 0 is the worst. The F1 score is obtained by calculating the precision and recall of the model. From Figure 10, out of 86 images it is seen that 80 are correctly predicted as having a diameter of 80mm and remaining 6 incorrectly assigned to other classes. So the precision of the diameter 80 mm class shown in Table 3, can be calculated as:

$$Precision = \frac{80}{86} = 0.93 \qquad (2)$$

In Figure 10, out of a total of 86 images actually labeled as 82 mm diameter, 80 images are correctly predicted and 2 of them are incorrectly predicted. Thus, the recall can be calculated as follows:

$$Recall = \frac{80}{82} = 0.98 \quad (3)$$

The F1-score can be calculated as follows: $F1 - score = 2 * \frac{Precision*recall}{Precision+recall} = 0.95 \quad (4)$

| Classification report | Precision | Recall | F1-score | Accuracy |
|---|---|---|---|---|
| Diameter 16 mm | 0.96 | 0.98 | 0.99 | |
| Diameter 24 mm | 1.00 | 0.98 | 0.99 | |
| Diameter 32 mm | 1.00 | 1.00 | 1.00 | |
| Diameter 40 mm | 1.00 | 0.95 | 0.98 | |
| Diameter 48 mm | 0.95 | 0.97 | 0.96 | |
| Diameter 64 mm | 0.85 | 0.98 | 0.96 | |
| Diameter 72 mm | 0.95 | 0.84 | 0.89 | 96% |
| Diameter 80 mm | 0.93 | 0.98 | 0.95 | |
| Diameter 100 mm | 1.00 | 0.98 | 0.99 | |

Table 3 : Results from the classification report: Diameter identification on test set.

The confusion matrix of the model is shown in Figure 10 . The y-axis corresponds to the actual class labels and the x-axis to the predicted class labels. The majority of diameters are well classified, except that the classifier produces some false negatives and false positives in the last three classes, which are 100mm, 72mm and 80mm.

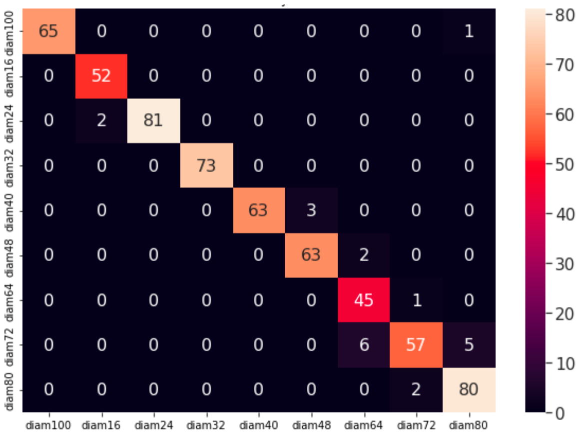


Figure 10: matrix-confusion of diameter identification - Resnet50 Model

Prediction on multiple buried objects: In order to evaluate the performance of the network in the case of various underground objects, a couple of B-Scans are created with up to 2 objects of different diameters and depths with distinct pipe types. The figure 11 illustrates a typical B-Scan image showing 2 hyperbolas with multiple interlocking borders. The hyperbola on the left corresponds to a metallic type pipe, diameter 40mm, and the hyperbola on the right corresponds to a water filled type pipe, diameter 32 mm, buried in a concrete medium. The hyperbolas were isolated by a-scan signal fitting processing. Subsequently, the extracted frames are reshaped to match the size of the initial dataset.

| Prediction | True labels | Predicted Labels |
|---|---|---|
| Diameter | 40mm | 40mm |
| Medium of propagation | Concrete | Concrete |
| Type of pipe | Metallic | Metallic |

Table 3: Prediction performance for multi-object scenarios

## 4 Conclusion

In this study, a multi-label deep learning model was proposed to identify different characteristics of buried objects from GPR B-scan signals. The three models developed in this work are: Resnet 50, VGG-16, and a custom CNN architecture, the main objective is to obtain a multiple identification of different characteristics of the objects: pipe type, diameter identification, as well as the propagation medium. In this study, a new CNN architecture is proposed while introducing the multi-path concept in order to extract the different characteristics at the same time. Future work will focus on generating noisy data to assess the robustness of our models. As well as the development of a new algorithm based on transformers for extracting characteristics from Radar signals while combining A-Scan and B-Scan.

## References and bibliography:

[1] T. Noreen and U. S. Khan, "Using pattern recognition with HOG to automatically detect reflection hyperbolas in ground penetrating radar data," *2017 International Conference on Electrical and Computing Technologies and Applications (ICECTA)*, Ras Al Khaimah, United Arab Emirates, 2017, pp. 1-6, doi: 10.1109/ICECTA.2017.8252064.

[2] E. Temlioğlu, M. Dağ and R. Gürcan, "Comparison of feature extraction methods for landmine detection using Ground Penetrating Radar," 2016 24th Signal Processing and Communication Application Conference (SIU), Zonguldak, Turkey, 2016, pp. 665-668, doi: 10.1109/SIU.2016.7495827.

[3] W. A. Wahab, J. Jaafar, I. M. Yassin and M. R. Ibrahim, "Interpretation of Ground Penetrating Radar (GPR) image for detecting and estimating buried pipes and cables," *2013 IEEE International Conference on Control System, Computing and Engineering*, Penang, Malaysia, 2013, pp. 361-364, doi: 10.1109/ICCSCE.2013.6719990.

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

[5] O. Missaoui, H. Frigui and P. Gader, "Model level fusion of edge histogram descriptors and gabor wavelets for landmine detection with ground penetrating radar," 2010 IEEE International Geoscience and Remote Sensing Symposium, Honolulu, HI, USA, 2010, pp. 3378-3381, doi: 10.1109/IGARSS.2010.5650350.

[6] Christian Maas and Jörg Schmalzl (2013): Using pattern recognition to automatically localize reflection hyperbolas in data from ground penetrating radar. In: Computers & geosciences, vol. 58,

[7] Iraklis Giannakis and Antonios Giannopoulos and Craig Warren (2020): A machine learning scheme for estimating the diameter of reinforcing bars using ground penetrating radar. In: IEEE Geoscience and Remote Sensing Letters, 2020.

[8] Barkataki, Nairit, Banty Tiru, and Utpal Sarma. "A CNN model for predicting size of buried objects from GPR B-Scans." Journal of Applied Geophysics 200 (2022): 104620

[9] V. Kafedziski, S. Pecov and D. Tanevski, "Detection and Classification of Land Mines from Ground Penetrating Radar Data Using Faster R-CNN," 2018 26th Telecommunications Forum (TELFOR), Belgrade, Serbia, 2018, pp. 1-4, doi: 10.1109/TELFOR.2018.8612117.

[10] S. Akçali and F. Erden, "Support of Data Augmentation with GAN on Faster R-CNN Based Buried Target Detection," 2021 29th Signal Processing and Communications Applications Conference (SIU), Istanbul, Turkey, 2021, pp. 1-4, doi: 10.1109/SIU53274.2021.9477828.

[11] V. Ciarletti, A. Herique, J. Lasue, R. Levasseur, C. Anny; D. Plettemeier, F. Lemmonier, C. Guiffaut, P. Pasquero, W. Kofman, “CONSERT constrains the internal structure of 67P at a few-meter size scale,” MNRAS, Monthly Notices Royal Astronomical Society, 469, 2018.

[12] Taflove, Allen, Susan C. Hagness, and Melinda Piket-May. "Computational electromagnetics: the finite-difference time-domain method." *The Electrical Engineering Handbook* 3 (2005): 629-670.

[13] Reddy, R. Pradeep Kumar, Chiluka Nagaraju, and I. Rajasekhar Reddy. "Canny scale edge detection." Int J Eng Trends Technol (IJETT). https://doi. org/10.14445/22315381/IJETT-ICGTETM N 3 (2016).