Multi-label classification with deep learning techniques applied to the Bscan images of radar GPR

Soukayna El Karakhi, Alain Reineix, Christophe Guiffaut

Multi-label classification with deep learning techniques applied to the Bscan images of radar GPR

Abouzouhour El karakhi Soukayna, Reineix Alain, Guiffaut Christophe
University of Limoges
XLIM Institute, 123 Av. Albert Thomas, 87000 Limoges, France
elkarakhi@xlim.fr

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 diffracted 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 image called B-scan. They show response shapes of hyperbolic type and their analysis give many characteristics. For example, in the case of buried pipes, a specific processing allows to find their diameter, their nature as well as the electrical 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 characterization 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 has been 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 a synthetic dataset created in this study and evaluated on a set of performance metrics.

Les radars à pénétration de sol (GPR) sont aujourd'hui 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 dans le domaine temporel consiste à émettre des impulsions électromagnétiques dans le sol, celles-ci étant ensuite diffractées par les cibles à détecter. Une seule trace de signal GPR capturée à une position du radar est un signal 1D appelé Ascan. Un ensemble de formes d'ondes radar Ascan capturées à un certain nombre de positions consécutives différentes le long d'une direction particulière formera une image 2D appelée B-scan dans le cas d'un déplacement rectiligne. Elles montrent 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 connaître leur diamètre, leur nature ainsi que les caractéristiques électriques du sol. Cependant, ces approches nécessitent souvent un post-traitement complexe du Bscan, ce qui peut être chronophage et rend donc difficile la caractérisation en temps réel au détriment de ces méthodes. Avec l'émergence des réseaux neuronaux profonds et avec une phase d'apprentissage sur un grand nombre de Bscan, il devient possible d'extraire presque 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 des éléments d'information multiples sur la même image et réaliser la 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 formés sur un ensemble de données synthétiques créé dans cette étude et évalués sur un ensemble de mesures de performance.

Keywords: Ground Penetrating Radar, Image processing, Detection of Buried objects, Deep learning.

1 Introduction

Ground Penetrating Radar (GPR) [Jol08] is an increasingly significant and effective approach for 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 [NK17] 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 [TDG16] 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 [Wah+13] 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 [WR19] 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 [MS13], Wavelet Edge Detection [GGW20], 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 [BTS22] [KPT18] [AE21] [Tis+09] increasingly began to employ CNN-based methods of learning hyperbolic shapes for the classifications and identification of buried objects. Most of the methods mentioned here focus on identifying and positioning of buried objects, and classifying them according to a single characteristic. Recently, very limited studies have focused on a multi-label classification model [TK07] [Wei+14] 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 mainly 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 [XLI] developed within the EMC team of the XLIM Institute. It consists of a FDTD-based method software [THP05] 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).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 [XLI], but also on homogeneous medium 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 4 mm in each cartesian direction.

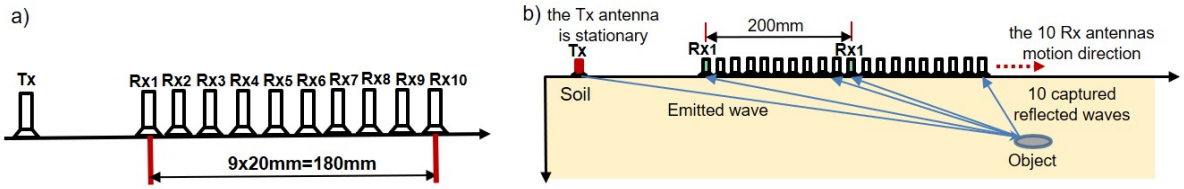

Fig 1: a) The multi-receiver 10 Rx antennas bistatic GPR with one Tx antenna emitter architecture, b) The way the 10 Rx antennas are moved in one direction during the exploration on the soil

The soil layer depth is 1 meter with a bandwidth of 500 MHz. 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. The waveform generated by the Wu-King antenna is a sinusoidal Gaussian, and this pulse is centered at the frequency of 1Ghz. This duration was calculated for the case of the concrete layer which has the smallest value for the propagation velocity among the media considered: $v_p = \frac{c}{\sqrt{\epsilon_r}}$ with c the velocity of light in vacuum. The GPR antennas system is composed of a fixed emitter and 10 receiving antennas. The receivers are also spaced 20 mm apart and moved in groups of 200 mm along the horizontal scan line on the soil, resulting in a 10-fold reduction in the number of simulations compared to considering the receivers individually throughout the computational volume. The multi-receiver bistatic GPR is illustrated in the Figure 1. 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.

2.1 A-Scan

An A-scan is a one-dimensional (1D) representation of the amplitude of the collected signal as a function of time after placing the receiving sensor antenna above the point of interest. An example of an A-scan is shown in Figure 2. It represents the signal captured by the receiving sensor antenna R_{x1} from the group of 10 low coupled receiving antennas. The first reflection generally corresponds to a combination of direct coupling between the antennas (Tx, R_{x1}) and reflection off the surface of the ground. The second reflection (at $t = 8.5$ ns), with a lower amplitude, corresponds to the response of the object of interest. To improve the quality of the recordings, it is often essential to eliminate both surface echoes and direct antenna coupling, retaining only the relevant signal. This operation can be performed using two methods during the simulation. Firstly, by taking a measurement without any objects present and subtracting it from the recording obtained with objects in the scene, effectively removing the various couplings. Alternatively, this can be achieved by disregarding the initial moments of the recording.

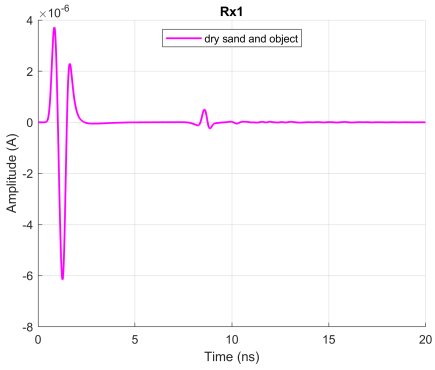

Fig 2: Example of the waveform of an A-scan captured by the receiving antenna. R_{x1}

2.2 *B-Scan*

The 2D image of the B-scan represents a vertical cross-sectional view of the subsurface. The horizontal axis of this section represents the position on the ground surface, while the vertical axis represents the round-trip time of the EM wave propagation. By combining all the A-scans that make up the B-scan, a matrix is created where each row represents a time sample and each column represents an A-scan trace. Each element of this matrix has a value corresponding to the amplitude of the field for the associated A-scan trace and time sample. Buried objects appear as hyperbolic shapes in the B-scan images. Recordings taken from different positions on the ground are combined in the B-scan image, thus creating these specific hyperbolic shapes. The figure 4 is illustrating the resulting treatment of B-scan.

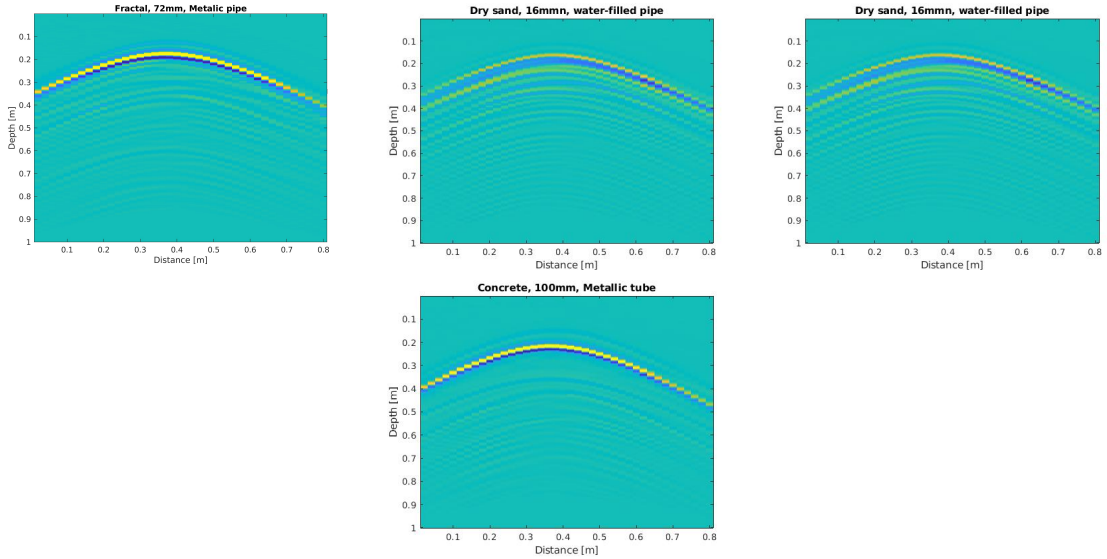

Figure: An example of B-scan images obtained using various propagation mediums, different diameters, and types of pipes.

2.3 *Relevant parameter values*

The parameter values employed in the simulation can be found in **Table 1**, while **Table 2** displays the dielectric parameters corresponding to the soil types.

Table 1: The parameter values of simulations

Parameters of simulations	**Values**
Source frequency	1.00 GHz
Bandwidth	500 MHz
Source waveform	Sino-Gaussian pulse
Source pulse duration	1 ns
A-scans intervals	20 mm
No of A-scans	40
Spatial resolution	4 mm

The table 2 provides information on different soil types along with their respective conductivity and relative permittivity values. The fractal model illustrates that soils are present in distinct layers or geological formations, which are separated by rough interfaces. These layers demonstrate a wide range of characteristics [Cia+17]

Table 2: Soil Conductivity and Relative Permittivity

SI. No	Soil type	Conductivity (S/m)	Relative permittivity (ε_r)
1	Dry sand	0.002	10.00
2	Dry clay	0.001	5.53
3	Concrete	0.001	8.00
4	Fractal	0.001	6.00

3 Proposed Deep Learning architecture

Now, the CNN and the procedure used to apply deep learning for retrieving the soil and pipes characteristics will be described. First, the global GPR data handling workflow in this study involves feature extraction, splitting the dataset into training and test sets. This allows respectively CNN model training and prediction efficiency. 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 [RNR16], which proceeds through the following steps:

1. noise reduction: Noise is removed from the image with a 5x5 Gaussian filter [BTS22].
2. Determine the gradient intensity of the image using the Canny operator.
3. Non-maximum Suppression: this step consists of eliminating those undesirable pixels that may not be edges.
4. Hysteresis Thresholding

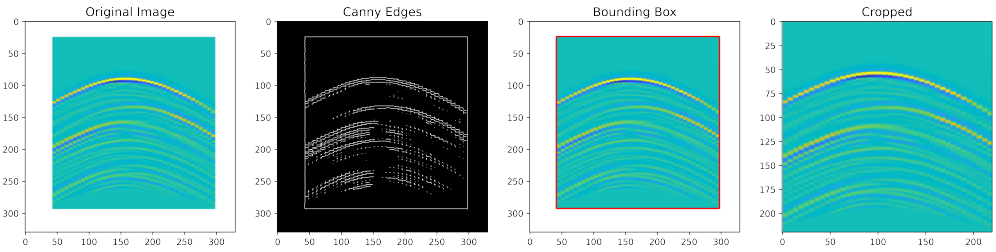

Fig 4: Example of output from edge detection operators

The objective of this work is to implement Deep Learning models for extracting 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, commonly known as ResNet, have revolutionized the field of computer vision by serving as a fundamental architecture for various tasks. The term "Residual" refers to the key concept behind ResNet's [Sar+21] success. Prior to ResNet, training deep neural networks with over 50 layers was challenging due to the problem of gradient backpropagation [TL19]. However, ResNet overcame this obstacle. The strength of ResNet lies in its implementation of skip connections, which are illustrated in Figure 5. 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. This connection, known as a skip connection, plays a crucial role in retaining important information from earlier layers. It enables the network to effectively train deep networks by mitigating the vanishing gradient problem.The skip connection in ResNet facilitates the flow of gradients throughout the network during training. This means that the gradients can propagate back to earlier layers more easily, allowing for smoother and more efficient learning. By incorporating these skip connections, ResNet has achieved remarkable performance in various computer vision tasks and has become a state-of-the-art model for deep neural network training.

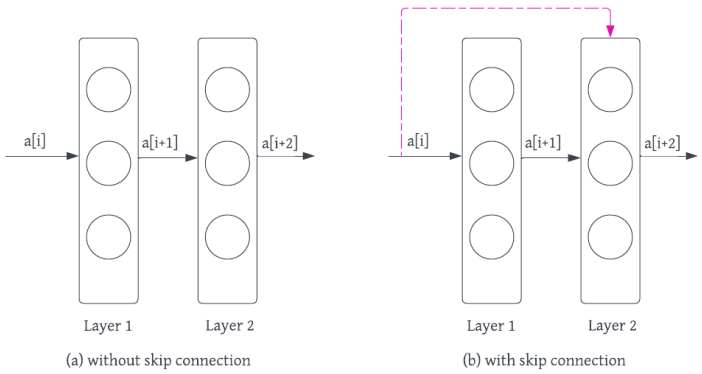

Fig 5: Residual learning: a building block

The Visual Geometry Group **(VGG-16)** [SZ14] is a widely recognized deep Convolutional Neural Network (CNN) architecture developed by the Visual Geometry Group. The key contribution of VGG-16 lies in its design, which utilizes a stack of 3x3 filters. This approach demonstrated

that using multiple small kernel size filters can achieve the same effect as using larger filters, while also providing the flexibility to increase the depth of the network. VGG models introduced the notion of depth as a crucial factor in the success of convolutional neural networks. By employing smaller filter sizes, VGG models are able of learning a larger number of filters, leading to a more complex representation of the input data. This approach marked a significant departure from earlier models that heavily relied on larger filters to capture low-level features.One of the notable achievements of VGG models is their ability to effectively capture both high-level and low-level features using smaller filters. This demonstrated that the depth of the network combined with smaller filters can produce powerful representations. This insight paved the way for subsequent advancements in deep learning architectures. In the case of VGG-16, the network comprises 16 trainable layers, as depicted in Figure 6. The VGG-16 model has been widely adopted in various computer vision tasks and has achieved state-of-the-art performance on benchmark datasets. Its success has inspired further research and the development of more advanced CNN architectures.

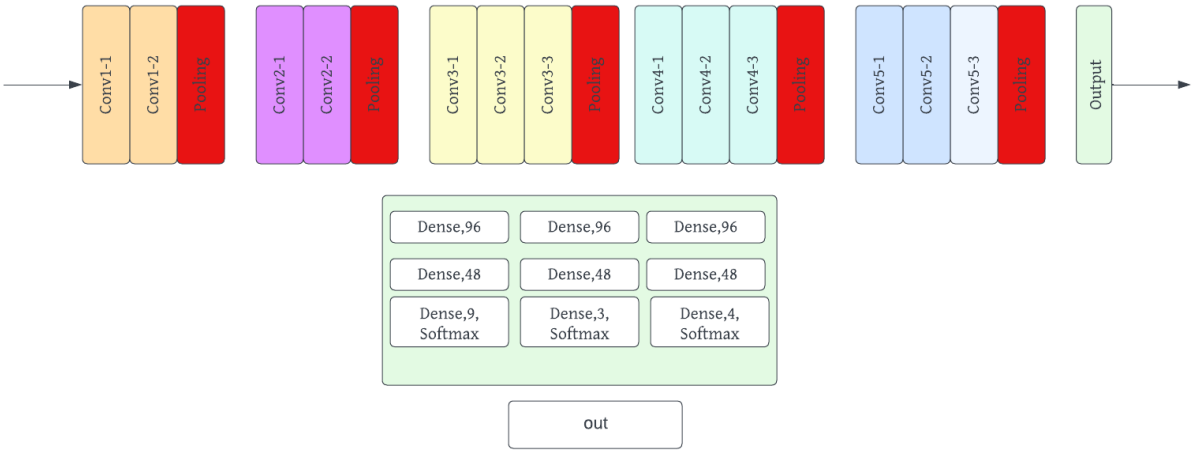

Fig 6: VGG16 architecture with building block of last layers

The pre-trained models VGG16 and Wide ResNet-50 are used in this study. Normalized input images in mini-batches, with RGB images having three channel dimensions (3×H×W), where H and W are assumed to be 224, are required by these models. In our approach, the final classification layers are substituted with fully connected layers. Specifically, three separate layers replace the last layer. Diameter classification, pipe type classification, and differentiation of propagation mediums are handled by the first, second, and third layers, respectively. ReLU and softmax activation functions are employed. The initial training layers are frozen, and the modified layer is trained using the B-Scans dataset. A training duration of 30 epochs with a batch size of 32 is used. Adam [Zha18] [TN18] serves as the model optimizer, and categorical cross-entropy is employed for loss calculation. To enhance computational efficiency and performance, techniques such as early stopping, data augmentation, and adaptive learning rates are employed. These techniques aim to improve both the speed of computation and the overall performance of the model.

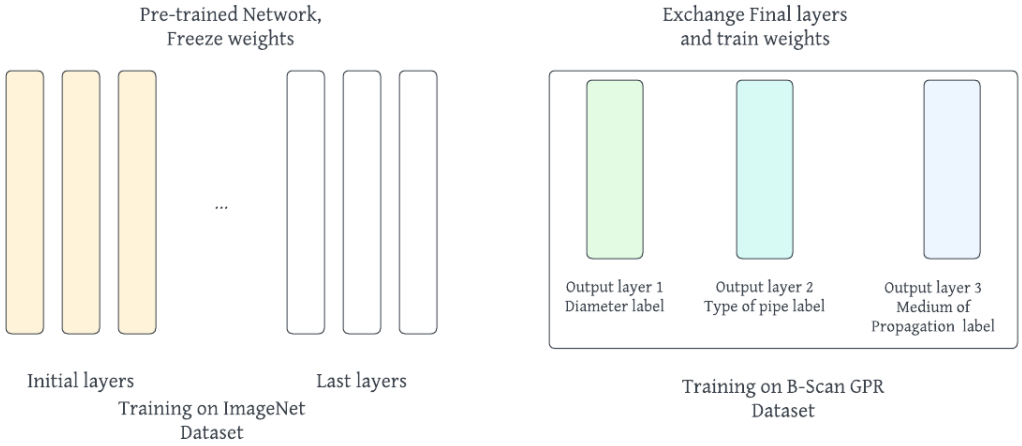

Fig 7: Transfer learning strategy

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 [Cao+18], it is a model developed by google. The main difference between the inception model and the ordinary CNNs [JZA18] 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. In the input layer, 32 filters of size 3×3 are used, followed by a MaxPool layer and then a batch normalization layer to increase the speed and stability of learning. After that, a convolution layer with 64 filters of size 3x3 is applied, followed by a max pooling layer. Two inception building blocks are then defined, each containing convolution layers with 128 and 16 filters of size (1x1), (5x5), (7x7), respectively. 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 [RP14] problems to predict the probability of each label.

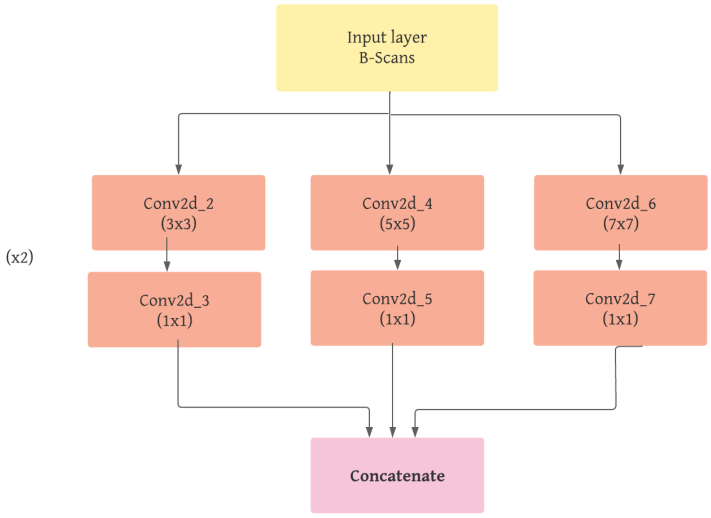

Fig 8: Custom CNN Model - building block

4 Results

The simulation dataset consists of nine different diameter values, three types of pipes, and four different propagation mediums. The dataset includes 50 different depths for each simulation scenario. Each simulation scenario generates a B-scan image, resulting in a total of 5,400 B-scan images. To further enhance the dataset, data augmentation techniques were applied, resulting in a total of 10,800 B-scan images. The data augmentation process involved key transformations, including no rotation of the images, zooming up to 10% of their original size, horizontal movement up to 10% of their width, and vertical movement up to 10% of their height.

4.1 Data augmentation

Enhancing model performance is a comprehensive process that involves employing various techniques to improve the accuracy, efficiency, and robustness of the model. This includes the use of advanced algorithms, preprocessing data, selecting relevant features, and tuning hyperparameters. By optimizing these factors, significant improvements can be achieved in the model's performance. The effectiveness and reliability of a deep learning model greatly depend on the quantity and diversity of the training data. To enhance data diversity and improve the model's interpretation of information, basic data augmentation techniques as described in reference [VM01] [SK19] [Hu+19] are employed. These techniques introduce visual variability to the data, leading to enhanced accuracy. For the B-Scans GPR dataset, the applied data augmentation techniques involve the use of Horizontal and Vertical shifts, as well as Resizing

4.2 Early stopping techniques

Early stopping [Bai+21] [Mah+17] is a regularization technique commonly used in deep neural networks. It involves monitoring the model's performance on a validation dataset during training and stopping the training process after a certain number of epochs if the model's performance does not improve. The weights of the model are saved and updated throughout training. When further updates no longer result in performance improvement, training is halted, and the most recent optimal parameters are kept.

4.3 Evaluation metrics

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:

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

Based on the results presented in Figure 9, a higher accuracy of 96% is observed for the diameter identification using the VGG16 mode, and an efficiency of 98% for the pipe type identification achieved by the Resnet-50 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.

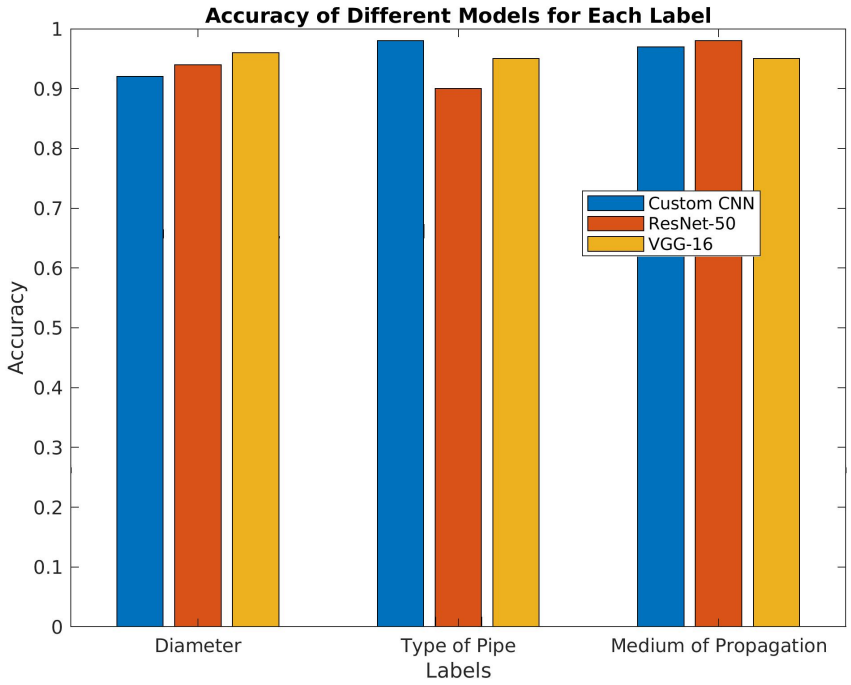

Fig 9: Accuracy vs. weighted accuracy of the three developed models with Canny's operators

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:

$$\text{Precision} = \frac{80}{86} = 0.93 \tag{2}$$

In Figure 10, out of a total of 86 images labeled as 80 mm diameter, 80 images were correctly predicted, resulting in a recall of approximately 93.00% for the class "diam80".

$$\text{Precision} = \frac{80}{82} = 0.98 \tag{3}$$

The F1-score can be calculated as follows:

$$\text{F1-score} = \frac{2 \cdot \text{Precision} \cdot \text{recall}}{\text{Precision} + \text{recall}} = 0.95 \tag{4}$$

Table 3: Results from the Classification Report: diameter Identification on Test Set

Labels	Precision	Recall	F1-Score
Diameter 16mm	0.96	0.98	0.99
Diameter 24mm	1.00	0.98	0.99
Diameter 32mm	1.00	0.98	1.00
Diameter 40mm	1.00	0.95	0.98
Diameter 48mm	0.95	0.97	0.96
Diameter 64mm	0.85	0.98	0.96
Diameter 72mm	0.95	0.84	0.89
Diameter 80mm	0.93	0.98	0.99
Diameter 100mm	1.00	0.98	0.99

In Figure 10, the confusion matrix provides insights into the model's accuracy in predicting diameters. It excelled in predicting diam16, diam32, and diam80, but misclassified diam72 as diam80 or diam64. The matrix depicts actual class labels on the vertical axis and predicted class labels on the horizontal axis. It showcases precise classification for all four categories (clay, concrete, fractal, sand) with no misclassifications. However, some instances of "Water-filled pipe" were misclassified as "Vaccum" while the model performed well in classifying "Metallic" and "Vaccum" categories.

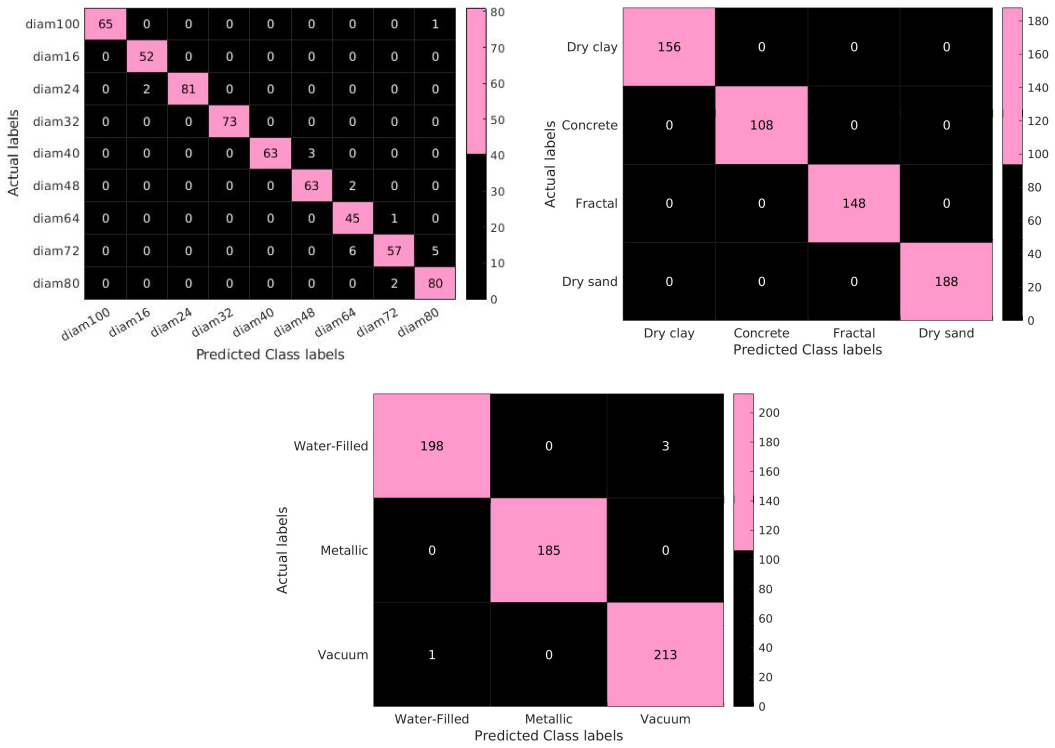

Figure: Matrix-confusion of diameter identification - Medium of propagation and type of pipe classification with Resnet-50 Model

4.4 Prediction on multiple buried objects

The network's performance was assessed by testing it on B-Scans containing multiple underground objects. These B-Scans were designed to have up to two objects with different diameters, depths, and pipe types. By evaluating the network's predictions on these complex scenarios, we gained insights into its ability to accurately identify and classify multiple objects simultaneously. The results showed a high accuracy of the model in successfully recognizing and distinguishing between the various underground objects, which demonstrates its effectiveness in handling real-world scenarios with multiple objects of diverse characteristics.

5 Discussion and conclusions

In this study, the focal point was on developing a multi-label classification (MLC) model using transfer learning and data augmentation techniques. Three deep learning models, namely VGG-16, ResNet-50, and an adapted CNN, were utilized as pre-trained models for transfer learning. These models were trained on a synthetic dataset created using the TEMSI-FD software, which simulates the propagation of electromagnetic waves in complex media. To enhance the model's performance, various techniques were employed, including advanced algorithms, data preprocessing, feature selection, and hyperparameter tuning. Data augmentation techniques, such as Horizontal and Vertical shift, as well as Resizing, were used to improve data diversity and the model's interpretation of information.The regularization technique of early stopping was applied to the training process. This technique involves monitoring the model's performance on a validation dataset and stopping the training after a certain number of epochs if no improvement is observed. The weights of the model

are saved and updated throughout training, and the most recent optimal parameters are used when further updates no longer improve performance. Based on the results presented the VGG16 model achieved a better accuracy of 96% in diameter identification, while the Resnet-50 model achieved an efficiency of 98% in pipe type identification. Both the VGG16 model and the customized CNN model achieved a higher level of sensitivity (98%) in the identification of the propagation medium. To evaluate the performance of the network in the presence of various underground objects, B-Scan images were created with up to 2 objects of different diameters, depths, and pipe types. In conclusion, this study successfully developed a multi-label classification model using deep learning techniques, transfer learning, and data augmentation. The models demonstrated high accuracy and sensitivity in identifying diameters, pipe types, and propagation mediums. The use of synthetic datasets, advanced algorithms, and regularization techniques contributed to the model's improved performance and reliability. 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

[VM01] David A Van Dyk and Xiao-Li Meng. "The art of data augmentation". In: *Journal of Computational and Graphical Statistics* 10.1 (2001), pp. 1–50.

[THP05] Allen Taflove, Susan C Hagness, and Melinda Piket-May. "Computational electromagnetics: the finite-difference time-domain method". In: *The Electrical Engineering Handbook* 3 (2005), pp. 629–670.

[TK07] Grigorios Tsoumakas and Ioannis Katakis. "Multi-label classification: An overview". In: *International Journal of Data Warehousing and Mining (IJDWM)* 3.3 (2007), pp. 1–13.

[Jol08] Harry M Jol. *Ground penetrating radar theory and applications*. elsevier, 2008.

[Tis+09] Bouchra Tissafi et al. "Comparative study of three wave propagation software programs for the modeling of coupled Maxwell and Boltzmann equations at THz frequency". In: *The Applied Computational Electromagnetics Society Journal (ACES)* (2009), pp. 382–390.

[MS13] Christian Maas and Jörg Schmalzl. "Using pattern recognition to automatically localize reflection hyperbolas in data from ground penetrating radar". In: *Computers & geosciences* 58 (2013), pp. 116–125.

[Wah+13] Warishah Abdul Wahab et al. "Interpretation of Ground Penetrating Radar (GPR) image for detecting and estimating buried pipes and cables". In: *2013 IEEE International Conference on Control System, Computing and Engineering*. IEEE. 2013, pp. 361–364.

[RP14] Jesse Read and Fernando Perez-Cruz. "Deep learning for multi-label classification". In: *arXiv preprint arXiv:1502.05988* (2014).

[SZ14] Karen Simonyan and Andrew Zisserman. "Very deep convolutional networks for large-scale image recognition". In: *arXiv preprint arXiv:1409.1556* (2014).

[Wei+14] Yunchao Wei et al. “CNN: Single-label to multi-label”. In: *arXiv preprint arXiv:1406.5726* (2014).

[RNR16] R Pradeep Kumar Reddy, Chiluka Nagaraju, and I Rajasekhar Reddy. “Canny scale edge detection”. In: *Int J Eng Trends Technol (IJETT). https://doi. org/10.14445/22315381/IJETT-ICGTETM N* 3 (2016).

[TDG16] Eyyup Temlioğlu, Mahmut Dağ, and Rıdvan Gürcan. “Comparison of feature extraction methods for landmine detection using ground penetrating radar”. In: *2016 24th Signal Processing and Communication Application Conference (SIU)*. IEEE. 2016, pp. 665–668.

[Cia+17] Valérie Ciarletti et al. “CONSERT constrains the internal structure of 67P at a few metres size scale”. In: *Monthly Notices of the Royal Astronomical Society* 469.Suppl_2 (2017), S805–S817.

[Mah+17] Maren Mahsereci et al. “Early stopping without a validation set”. In: *arXiv preprint arXiv:1703.09580* (2017).

[NK17] T Noreen and Umar S Khan. “Using pattern recognition with HOG to automatically detect reflection hyperbolas in ground penetrating radar data”. In: *2017 International Conference on Electrical and Computing Technologies and Applications (ICECTA)*. IEEE. 2017, pp. 1–6.

[Cao+18] Yang-Jie Cao et al. “Recent advances of generative adversarial networks in computer vision”. In: *IEEE Access* 7 (2018), pp. 14985–15006.

[JZA18] Nadia Jmour, Sehla Zayen, and Afef Abdelkrim. “Convolutional neural networks for image classification”. In: *2018 international conference on advanced systems and electric technologies (IC_ASET)*. IEEE. 2018, pp. 397–402.

[KPT18] V Kafedziski, S Pecov, and D Tanevski. *Detection and Classification of Land Mines from Ground Penetrating Radar Data Using Faster R-CNN 26th Telecommunications Forum (TELFOR)*. 2018.

[TN18] Ange Tato and Roger Nkambou. “Improving adam optimizer”. In: (2018).

[Zha18] Zijun Zhang. “Improved adam optimizer for deep neural networks”. In: *2018 IEEE/ACM 26th international symposium on quality of service (IWQoS)*. Ieee. 2018, pp. 1–2.

[Hu+19] Benlin Hu et al. “A preliminary study on data augmentation of deep learning for image classification”. In: *arXiv preprint arXiv:1906.11887* (2019).

[SK19] Connor Shorten and Taghi M Khoshgoftaar. “A survey on image data augmentation for deep learning”. In: *Journal of big data* 6.1 (2019), pp. 1–48.

[TL19] Hong Hui Tan and King Hann Lim. “Vanishing gradient mitigation with deep learning neural network optimization”. In: *2019 7th international conference on smart computing & communications (ICSCC)*. IEEE. 2019, pp. 1–4.

[WR19] Benjamin Walker and Laura 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*. IEEE. 2019, pp. 3578–3581.

[GGW20] Iraklis Giannakis, Antonios Giannopoulos, and Craig Warren. “A machine learning scheme for estimating the diameter of reinforcing bars using ground penetrating radar”. In: *IEEE Geoscience and Remote Sensing Letters* 18.3 (2020), pp. 461–465.

[AE21] Sertan Akçali and Fatih Erden. “Support of Data Augmentation with GAN on Faster R-CNN Based Buried Target Detection”. In: *2021 29th Signal Processing and Communications Applications Conference (SIU)*. IEEE. 2021, pp. 1–4.

[Bai+21] Yingbin Bai et al. “Understanding and improving early stopping for learning with noisy labels”. In: *Advances in Neural Information Processing Systems* 34 (2021), pp. 24392–24403.

[Sar+21] Devvi Sarwinda et al. “Deep learning in image classification using residual network (ResNet) variants for detection of colorectal cancer”. In: *Procedia Computer Science* 179 (2021), pp. 423–431.

[BTS22] Nairit Barkataki, Banty Tiru, and Utpal Sarma. “A CNN model for predicting size of buried objects from GPR B-Scans”. In: *Journal of Applied Geophysics* 200 (2022), p. 104620.

[XLI] Software Developed in the XLIM Institute. *Time Electromagnetic Simulator - Finite Difference Time Domain*. Limoges, France.