

## A DEEP LEARNING APPROACH FOR URBAN UNDERGROUND OBJECTS DETECTION FROM VEHICLE-BORNE GROUND PENETRATING RADAR DATA IN REAL-TIME

Z. Zong<sup>1</sup>, C. Chen<sup>1,\*</sup>, X. Mi<sup>1</sup>, W. Sun<sup>1</sup>, Y. Song<sup>1</sup>, J. Li<sup>1</sup>, Z. Dong<sup>1</sup>, R. Huang<sup>2</sup>, B. Yang<sup>1,\*</sup>

<sup>1</sup> State Key Laboratory of Information Engineering in Surveying, Mapping and Remote Sensing, Wuhan University, Wuhan 430079, Hubei Province, PR China - (zlong, chichen, mixiaoxin, 2015301610165, cying\_syh, lijianping, 00032011, bsyang)@whu.edu.cn

<sup>2</sup> State Key Laboratory of Geodesy and Earth's Dynamics, Institute of Geodesy and Geophysics of CAS, Wuhan 430077, Hubei Province, PR China - gang3217@asch.whigg.ac.cn

### ICWG II/III: Pattern Analysis in Remote Sensing

**KEY WORDS:** Ground Penetrating Radar, Object Detection, Convolutional Neural Network, Deep Learning, Urban Space Security

#### ABSTRACT:

GPRs (Ground Penetrating Radar) are widely adopted in underground space survey and mapping, because of their advantages of fast data acquisition, convenience, high imaging resolution and NDT (Non Destructive Testing) inspection. However, at present, the automation of the GPR data post-processing is low and the identification of underground objects needs expert interpretation. The heavy manual interpretation labor limits the GPR applications in large-scale urban scenarios. According to the latest research, it is still an unsolved problem to detect targets or defects in GPR data automatically and needs further exploration. In this paper, we propose a deep learning method for real-time detection of underground targets from GPR data. Seven typical targets in urban underground space are identified and labelled to construct the training dataset. The constructed dataset is consist of 489 labelled samples including rainwater wells, cables, metal/nonmetal pipes, sparse/dense steel reinforcement, voids. The training dataset is further augmented to produce more samples. DarkNet53 convolutional neural network (CNN) is trained using the constructed training dataset including realistic data and augmented data to extract features of the buried objects. And then the end-to-end YOLO detection framework is used to classify and locate the seven specific categories buried targets in the GPR data in real time. Experiments show that the automatic real-time detection method proposed in this paper can effectively detect the buried objects in the ground penetrating radar image in real time at Shenzhen test site (typical urban road scene).

### 1. INTRODUCTION

With the rapid process of urbanization, the development and utilization of urban underground space has aroused great attention. Accurate and efficient detection of urban underground space targets, as well as identifying their types and distribution conditions are the premise and key to ensure the safety of urban underground space. The targets underground the city include a large number of artificial structures (e.g. subway, pipeline corridor) and complex potential diseases (e.g. empty holes). Traditional surveying and mapping methods based on photography and laser scanning are not penetrating, and cannot effectively detect the positions and characteristics of urban space targets underground. Ground Penetrating Radar (GPR) has the advantages of high efficiency, non-destructive, penetrability and high imaging resolution, making it an essential role from geophysical exploration (Enze Z, 2017), building quality detection, road and bridge detection, tunnel quality detection to underground detection and classification (LAI W L and DÉROBERT X, 2017). Because of the complex structures of underground target and the diversity of potential diseases, the analysis of GPR image still relies mainly on human-computer interaction to locate and detect underground targets, which cannot meet the needs of large-scale urban underground space exploration and census.

The main detection methods for underground pipeline targets in GPR images are hyperbolic feature extraction based on Hough transform (Windsor C and Capineri L, 2014; Li W and Cui X,

2016), which is limited by the huge amount of computation caused by processing and discretizing a large number of parameters. The hyperbolic feature based on template matching requires manual design of a large number of parameters to depict different target features (Sagnard F and Tarel J, 2016; Terrasse G and Nicolas J, 2016). The automation and accuracy of feature-based gradient direction histogram (HOG) and Haar-like features algorithm for underground target detection in GPR images need to be further improved to eliminate false alarm and missed detection (Torrione P and Morton K, 2014; Maas C and Schmalzl J, 2013).

In recent years, breakthroughs have been made in optical image target detection based on deep learning (Krizhevsky A and Sutskever I, 2012; Girshick R, 2015). In underground target detection with hyperbolic echo characteristics in GPR images, some scholars proposed a deep learning method based on FASTER RCNN (Pham M and Lefèvre S, 2018). However, due to the variety and complexity of underground targets, it is impossible to detect underground space targets precisely and effectively only depending on one feature. In order to solve the problems above, this paper proposed a real-time deep learning method for underground target detection based on YOLO V3 target detection and Darknet-53 convolution neural network (REDMON J and FARHADI A, 2018), constructed several typical underground target databases, and verified the method by taking the vehicle-borne GPR image of Shenzhen urban road as an experiment. The acceptable result shows the average precision and recall rate of target detection, both of the two are over 85%.

\* Corresponding author

## 2. DETECTION OF UNDERGROUND TARGETS IN GPR IMAGES BASED ON DEEP LEARNING

The flow chart of GPR image underground target detection method based on deep learning presented in this paper is shown in Figure 1. The three key steps of this method are as follows:

1. Constructing the sample data set of underground targets. It mainly completed the labeling of underground target samples and data compatibility through human-computer interaction, and constructed underground targets standard data sets.

2. Using joint training mechanism to train convolutional neural network. Darknet-53 convolutional neural network was trained with ImageNet data set, COCO data set and PASCAL VOC data set, through which the set of pre-trained neuron parameters were obtained.

3. Training and inference. Using transfer learning method, firstly, the first 50 layers of network parameters were frozen. Then the first 50 layers of parameter sets were unfrozen to join the training process and obtains the final model which can predict the location of multiple boundary boxes and multiple types in real time. Finally, the model after network training is loaded, inference is used to obtain the location and category information of underground targets in the test data set.

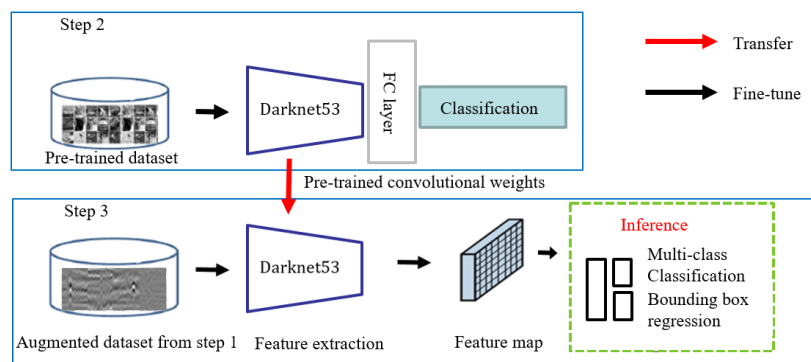


Figure 1. Target Detection from GPR imagery based on Convolutional Neural Network

### 2.1 Construction of Underground Target Training Samples

The purpose of the initial phase is to obtain enough labelled data for training the CNN (Convolutional Neural Network). We used the SIR-30 vehicle-borne GPR system to collect data from a typical city area at a frequency of 400 Hz. And manually labelled targets in the data, including rainwater wells, cables, metal/nonmetal pipes, sparse/dense steel reinforcement. The albumations library (Alexander B and Alex P, 2018) is used to augment the actual GPR data collected. Considering the resolution of the GPR image and the characteristic of the underground target, the GPR image data augmentation involves a combination of random cropping, small angle rotation, blurring, mirror flip, etc.

### 2.2 Underground Target Detection Network

Darknet-53 takes both network complexity and detection accuracy into account. Compared with VGG-16 (Simonyan K, 2014), it has less computation, lighter model and stronger feature extraction ability. It consists of consecutive 3\*3 and 1\*1 convolution kernel, which reduces the amount of network parameters in case of fully extracting image features. Drawing on ResNet idea, residual blocks are added to solve the problem of gradient dispersion and/or explosion caused by the deeper layers of the network. And the traditional deep neural network training is adjusted to stage-by-stage from layer-by-layer. In addition, Batch\_Normalize (BN) layer (IOFFE S, 2015) and LeakyReLU layer (Clevert D and Unterthiner T, 2016) are joined after each convolution layer, which can simplify the calculation while accelerating convergence and also prevent over-fitting.

### 2.3 Training and Inference

The network training method in this paper followed the network training method proposed by YOLO V3. The anchor box obtained by K-means clustering was used to assist the prediction of the boundary box, and the logistic regression classifier was trained to predict the object score of each boundary box. Each bounding box can contain multiple predicted categories. The neural network directly convoluted the GPR image to form a feature map, and then predicted the location and probability of the underground target grid by grid. The core of the algorithm was to transform the underground target detection problem into a regression problem, and realized the end-to-end real-time detection.

After the GPR image was input, the feature map of N\*N was formed by CNN convolution. Since YOLO V3 is a multi-layer detection, the values of N are 13, 26 and 52, respectively. If the centre of the target fall in a grid cell, the grid cell assumes responsibility for the inference of that target. In this method, three boundary boxes are predicted for each grid element. Same as the original frame, each boundary box output 12 prediction results, including x, y, w, h, confidence and seven conditional class probabilities. (x, y, w, h) is the absolute coordinate calculated from the original prediction results. Among them, x, y represent the value of the centre of the boundary box relative to the image boundary. And w, h are the length and width of the boundary box. The formula for calculating confidence is as follows:

$$Conf(Object) = Pr(Object) * IOU_{pred}^{truth} \quad (1)$$

If the centre of the ground truth does not fall in the cell, the confidence score is zero. Otherwise, the confidence score is equal to the Intersection over Union (IOU) between the prediction box

and the ground truth. In addition, each grid predicts seven conditional class probabilities, expressed as  $Pr(Class|Object)$ , which are conditioned on the inclusion of objects in grid cell. The formula for calculating the confidence of a particular category of each box is as follows:

$$Conf(Class|Object) = Pr(Class|Object) * Pr(Object) * IOU_{pred}^{truth} \quad (2)$$

In this paper, the sum-squared errors between output and actual vectors with both  $N*N*(2*5+7)$  dimension are used as loss function optimization parameters. In order to enhance the detection of small targets,  $\lambda_{coord} = (2 - truth_w * truth_h)$  was introduced to correct the coordinate errors, while  $\lambda_{obj}$  and  $\lambda_{class}$  were set to 1, so that the model can converge in training. The final loss function is:

$$Loss = \sum_{i=0}^W \sum_{j=0}^H \sum_{k=0}^A g(maxIOU) * \lambda_{noobj} * (-b_{ijk}^0)^2 + f(maxIOU)(\lambda_{coord} \sum_{r \in (x,y,w,h)} (truth_k^r - b_{ijk}^r)^2 + \lambda_{obj} * (1 - b_{ijk}^0)^2 + \lambda_{class} * (\sum_{c=1}^C (truth^c - b_{ijk}^c)^2)) \quad (3)$$

$$g(x) = \begin{cases} 1, & x < Thresh \\ 0, & x \geq Thresh \end{cases} \quad (4)$$

$$f(x) = \begin{cases} 0, & x < Thresh \\ 1, & x \geq Thresh \end{cases} \quad (5)$$

The loss function is mainly divided into three parts: coordinate loss, confidence loss and classification loss.  $maxIOU$  denotes the maximum IOU between the prediction box and the ground truth of a specific cell;  $Thresh$  is a pre-set threshold of 0.6;  $b_{ijk}^0$  is the confidence of predicted box;  $truth_k^r$  and  $b_{ijk}^r$  represent the coordinates of the real box and the predicted box respectively; while  $truth^c$  and  $b_{ijk}^c$  represent the classification of the real box and the predicted box respectively. Loss value is still calculated using SSE (The sum of squares due to error) in the same way as YOLO V1.

In YOLO V3 algorithm design, multi-label classification is used to enhance the ability of boundary box to contain multi-categories. Usually, a target only falls in one grid element. However, under some circumstances targets can be predicted by multiple grid cells, such as the prediction of large targets or the target near the boundary of several grid. YOLO algorithm before YOLO V3 can only predict a boundary box for each target, so it is very likely to miss detection. Unlike YOLO9000 (Redmon J and Farhadi A, 2016), YOLO V3 abandons the softmax function,

which has some limitations in improving the performance of network detection for multi-target. Instead, it applies a separate logical classifier to predict multi-label. In the training process, the binary cross-entropy loss is used to calculate the category loss.

The method in this paper uses Adam (Diederik P and Jimmy B, 2014) instead of the traditional Stochastic Gradient Descent (SGD) optimization algorithm. Similar to other optimization algorithms of deep learning, the weights of neural networks are updated iteratively after back propagation based on training data. To be more precise, it is a method of learning rate self-adaption. The Figure 2 shows the line chart of loss values varying with the number of epoch.

As can be seen from loss curve, the loss value does not change after about 85 epochs, and finally a stable model is obtained. The loss is stable at about 5. To sum up, underground target detection is to extract the feature map of the input GPR image, and then determine whether there are underground targets in it grid-by-grid. Similar to the forward propagation in the network training process, only one prediction is needed to obtain the boundary boxes of the possible targets, the target categories and their probabilities of each box, and then the prediction results are refined according to the pre-set thresholds. According to the actual situation, the IOU threshold is set to 0.3 and the confidence threshold is set to 0.5. Pulse answer that does not belong to any of the mentioned classes are mostly filtered because the confidence is below 0.5.

### 3. EXPERIMENTS AND ANALYSIS

In order to verify the effectiveness of this method, SIR-30 vehicle-based GPR system was used to test the GPR data of a go-return route in Caitian Road, Shenzhen City. The parameters such as antenna centre frequency, acquisition length and scene category are shown in Table 1. Table 2 shows the number of samples interpreted by experts and the number of samples after expansion. For the purpose of ensuring the convergence of the model and not over-fitting, the number of data after expansion is an empirical value in a suitable range.

#### 3.1 Detection Results of Underground Targets

The test data in this paper are the data of Shenzhen City obtained by SIR-30 vehicle-based GPR system at 400 MHz frequency. Figure 3(a)-(e) are the recognition results of rainwater wells, sparse/dense steel mesh, bridges, metal/non-metal pipelines, cables in GPR waveform image data, in which bridges are represented with two parts: dense steel mesh and empty. The recognition results show that the deep learning method proposed in this paper can accurately detect and locate the type and position of underground targets according to GPR images.

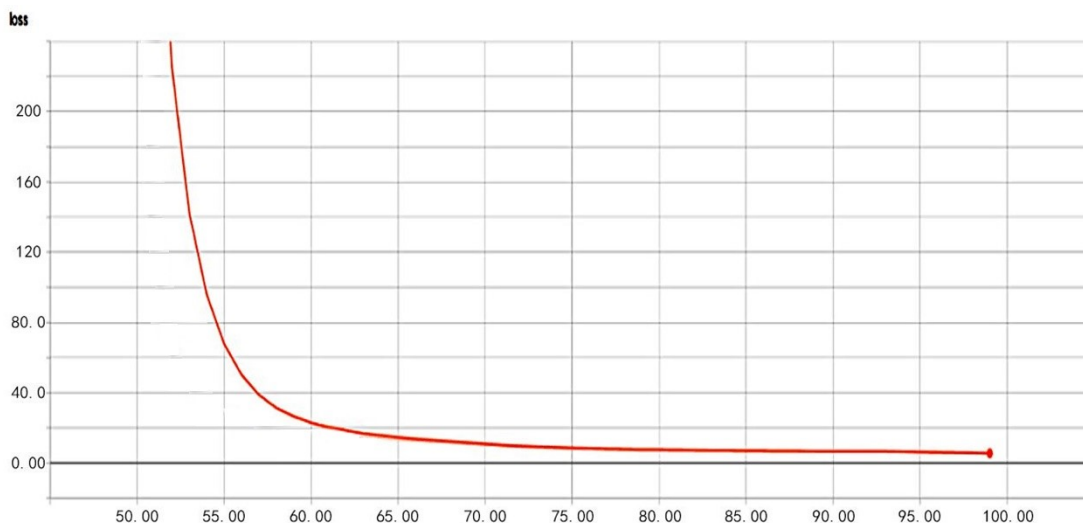


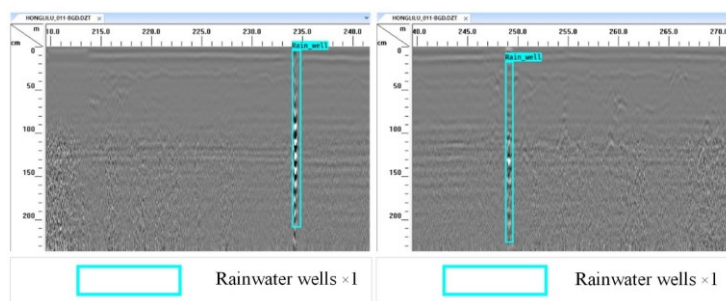
Figure 2. The line chart of Loss value

| Instrument model | Antenna center frequency | Test Area              | Scene category | Acquisition length | Scanning method     |
|------------------|--------------------------|------------------------|----------------|--------------------|---------------------|
| SIR-30           | 400MHz                   | Caitian Road, Shenzhen | City           | 4.52km             | Continuous scanning |
| SIR-30           | 400MHz                   | Caitian Road, Shenzhen | City           | 4.66km             | Continuous scanning |

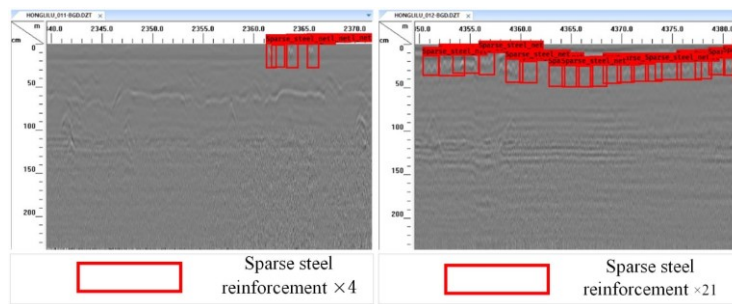
Table.1 GPR equipment and test area

| Classes                        | Quantity       |           |
|--------------------------------|----------------|-----------|
|                                | Interpretation | Augmented |
| Rainwater wells                | 57             | 570       |
| Cables                         | 32             | 581       |
| Metal pipes                    | 42             | 571       |
| Non-metal pipes                | 17             | 246       |
| Sparse steel reinforcement net | 87             | 264       |
| Dense steel reinforcement net  | 92             | 307       |
| Voids                          | 162            | 494       |
| Total                          | 489            | 3033      |

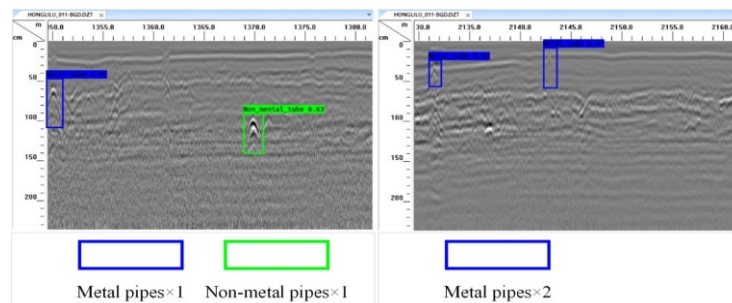
Table.2 Statistical quantitative table of samples of different categories after expansion



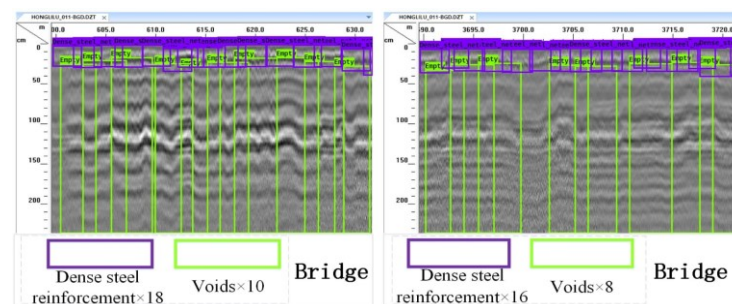
(a) Recognition results of rainwater wells



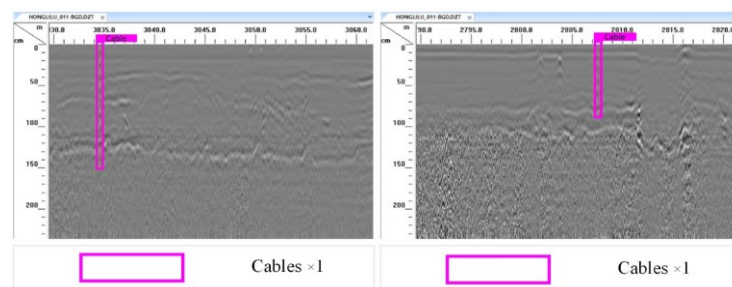
(b) Recognition results of sparse steel reinforcement



(c) Recognition results of metal pipes and non-metal pipes



(d) Recognition results of bridges



(e) Recognition results of cables

Figure 3. Recognition results

### 3.2 Evaluation of Detection Results

In the classification tasks of underground target detection, three indicators, precision and recall, and F1 score usually measure the performance of the network model. F1 score is calculated to evaluate precision and recall comprehensively. The precision rate represents the proportion of the classified objects correctly identified to the total targets, which reflects the ability of the model to distinguish the background in the image. The recall rate represents the proportion of a certain category of objectives eventually classified into that category, which indicates retrieval ability of the model for the targets in GPR images. The higher

precision indicates that the detection model can better distinguish the target from the background, whereas the higher the recall rate depicts the stronger the detection ability of the detection model for underground targets. F1 Score is a comprehensive expression of both sides. A high score can present a more robust classification model. The formulas for the three indicators are as follows:

$$\text{Precision} = \frac{TP}{TP + FN} \quad (6)$$

$$Recall = \frac{TP}{TP + FN} \quad (7)$$

$$F1 = \frac{2 * Recall * Precision}{Recall + Precision} \quad (8)$$

Among them, TP denotes the number of targets detected correctly, FP is the number of targets not detected, and FN represents the number of targets detected incorrectly. Table 3 shows the evaluation results of target detection in GPR images under a long-distance road in Caitian experimental area of Shenzhen.

| Classes                    | Metrics |           |         |
|----------------------------|---------|-----------|---------|
|                            | Recall  | Precision | F1score |
| Rainwater wells            | 0.912   | 0.954     | 0.932   |
| Cables                     | 0.923   | 0.857     | 0.889   |
| Metal pipes                | 0.852   | 0.920     | 0.885   |
| Non-metal pipes            | 0.833   | 0.833     | 0.833   |
| Sparse steel reinforcement | 0.942   | 0.867     | 0.903   |
| Dense steel reinforcement  | 0.969   | 0.839     | 0.900   |
| Voids                      | 0.816   | 0.816     | 0.816   |
| Average                    | 0.892   | 0.869     | 0.880   |

Table 3. Evaluation of the GPR image data target detection result

The experimental results show that all three evaluate indices, the average values of recall, precision and F1 Score of the deep learning method designed in this paper are over 85%. The comprehensive evaluation of the three indicators indicates that the method here is very effective for the detection of underground targets and defects in GPR images.

#### 4. CONCLUSION

In order to solve the problem of efficient and accurate detection of underground targets in GPR images, this paper proposed and designed a method of detecting underground targets in GPR images based on deep learning. Through data compatibility, the problem of insufficient underground target samples in GPR image was solved. Darknet 53 network drawn lessons from ResNet idea to add residual blocks, which can avoid gradient dispersion or explosion even if the network level was very deep. What's more, a large number of open data sets were used to pre-train the network, ensuring the ability of the network to extract the features of underground targets after training. And the multi-branch prediction in YOLO V3 framework also guaranteed the detection of small underground targets.

The actual scene test results show that this method can detect and classify underground targets in real time from vehicle GPR images. The detection speed can reach 16 frames/s on two GTX 1080 GPU platforms, and the average precision and recall rate can reach more than 85%. At present, the method in this paper only detects the category and location of the targets. Future research will combine the structural information of underground targets, including buried depth, area and aspect ratio, to construct feature space and random forest algorithm will be used to extract these features to assist in identifying target categories and improving the detection quality.

#### ACKNOWLEDGEMENTS

The work presented in this article was substantially supported by National Natural Science Foundation Project of China (No. 41701530, 41725005 and 41531177), China Postdoctoral

Science Found (No. 2018T110802), China Southern Power Grid Corporation Science and Technology Project (ZBKJXM20170229) and LIESMARS Special Research Funding and Key Laboratory of Spatial Data Mining & Information Sharing of Ministry of Education, Fuzhou University (No. 2018LSDMIS06).

#### REFERENCES

Alexander, B., Alex, P., Eugene, K., 2018. Albumentations: fast and flexible image augmentations. arXiv:1809.06839(18 Sep 2018).

Clevert, D., Unterthiner, T., Hochreiter, S., 2016. Fast and Accurate Deep Network Learning by Exponential Linear Units (ELUs). *Computer Science*. arXiv:1511.07289(22 Feb 2016).

Diederik, P., Jimmy, B., 2017: Adam: A Method for Stochastic Optimization. arXiv:1412.6980v9.(30 Jan 2017).

Enze, Z., 2017. Ground penetrating radar (GPR) application in urban engineering survey research. Anhui University of Science and Technology.

Girshick, R., 2015. Fast R-CNN. IEEE International Conference on Computer Vision. IEEE Computer Society, 2015, 1440-1448.

Ioffe, S., 2015. Batch Normalization: Accelerating Deep Network Training by Reducing Internal Covariate Shift. arXiv:1502.03167(2 Mar 2015).

Krizhevsky, A., Sutskever, I., Hinton, G., 2012. ImageNet classification with deep convolutional neural networks. *Advances in neural information processing systems*, 1097-1105.

Lai, W., Dérobert, X., Annan P., 2017. A review of Ground Penetrating Radar application in civil engineering: A 30-year journey from Locating and Testing to Imaging and Diagnosis. *Ndt & E International*, 58-78.

Li, W., Cui, X., Guo, L., Chen, J., 2016. Tree root automatic recognition in ground penetrating radar profiles based on randomized Hough transform. *Remote Sens.*, 430. doi.org/10.3390/rs8050430.

Maas, C., Schmalzl, J., 2013. Using pattern recognition to automatically localize reflection hyperbolas in data from ground penetrating radar. *Computers & Geosciences*, 58, 116–125.

Pham, M., Lefèvre, S., 2018. Buried object detection from B-scan ground penetrating radar data using faster-rcnn. arXiv:1803.08414v1(22 Mar 2018).

Redmon, J., Farhadi A., 2016: YOLO9000: Better, Faster, Stronger. arXiv:1612.082(25 Dec 2016).

Redmon, J., Farhadi, A., 2018. YOLOv3: An Incremental Improvement. arXiv:1804.02767(8 Apr 2018).

Sagnard, F., Tarel, J., 2016. Template-matching based detection of hyperbolas in ground-penetrating radargrams for buried utilities. *Geophys. Eng.*,13(4), 491.

Simonyan, K., 2014. Very deep convolutional networks for large scale Image recognition. *Computer Vision and Pattern Recognition*, 1-14.

- Terrasse, G., Nicolas, J., Trouv'e, E., Drouet, E., 2016. Automatic localization of gas pipes from gpr imagery. in *EUSIPCO*, 2395–2399.
- Torrione, P., Morton, K., Sakaguchi, R., Collins, L., 2014: Histograms of oriented gradients for landmine detection in ground-penetrating radar data. *IEEE Trans. Geosci. Remote Sens.*, 52(3), 1539–1550.
- Windsor, C., Capineri L., Falorni P., 2014. A data pair-labeled generalized Hough transform for radar location of buried objects. *IEEE Geosci. Remote Sens. Lett.*, 124–127. doi.org/10.1109/LGRS.2013.2248119.