A Study of Automatic Recognition and Localization of Pipeline for Ground Penetrating Radar Based on Deep Learning

Haobang Hu, Hongyuan Fang, Niannian Wang, Hai Liu, Jianwei Lei, Duo Ma, and Jiaxiu Dong

***Abstract*— This letter proposes a method based on deep learning for the automatic recognition and localization of underground pipelines using the ground penetrating radar (GPR). First, an automatic recognition model with an average precision (AP) of 0.9256 is proposed and trained based on Faster R-CNN. The feature extraction is optimized by the Attention-guided Context Feature Pyramid Network (ACFPN), and the cascade structure is used to improve the detection frame regression accuracy. Moreover, using Tesseract OCR, a positioning model is developed based on recognition results to obtain the burial and horizontal position of the pipeline. Furthermore, on-site experiments were carried out on real embedded pipes to verify the feasibility and effectiveness of the developed method. The absolute error of the localization data is lower than 11 cm, and the average error ratio is smaller than 12%. Consequently, it is demonstrated that the proposed method is considerably automatic, efficient, and reliable for the recognition and localization of underground pipelines.**

***Index Terms*— Faster R-CNN, ground penetrating radar (GPR), pipeline, recognition and localization.**

I. Introduction

THE ground penetrating radar (GPR) method identifies the underground target according to the discontinuity of the underground medium [1]. The GPR enjoys the advantages of being highly fast, highly efficient, continuous, and nondestructive. Two main target features of GPR images are hyperbolic morphological features and linear morphological ones, and the former has received more attention for pipeline detection. First, the visual interpretation of GPR data by an experienced practitioner is extremely time-consuming and labor-intensive [2]. Moreover, at present, the recognition method chiefly has some shortages of overreliance on manual subjective experience, slow speed, and low degree of automation. To overcome these limitations, the research on the GPR has gradually been switching to more effective data processing algorithms and object recognition methods [3], [4], [5], [6], [7]. The methods of hyperbolic feature recognition of pipelines largely include the traditional method, the machine learning (ML) method, and the deep learning method.

Deep learning skips the step of "feature engineering" and can learn the feature representation of buried objects directly from the GPR B-scan data. Besaw and Stimac [8] and Krizhevsky *et al.* [9] used convolutional neural networks (CNNs) to extract meaningful features from GPR B-scan data and classify buried explosives. Later, data augmentation (DA) was used to increase the number and variability of GPR training data, which markedly enhanced the performance of CNNs and traditional feature extraction algorithms. This method was extended to classify antitank mines and anti-infantry mines, improving the classification accuracy. Dinh *et al.* [10] also integrated the conventional image processing methods such as migration, filtering, threshold segmentation, and deep CNNs to extract a target region and send it into the trained CNNs; they then employed the trained model to test the dataset. In another work, Reichman *et al.* [11] introduced pretraining initialization steps and DA skills into the framework and discussed three different frameworks of CNNs, which was helpful in achieving higher performance of target detection. Furthermore, Pham and Lefèvre [12] used Faster R-CNN to detect reflective hyperbolas in GPR B-scans. Using gprMax [13] to simulate GPR B-scans, a CNN was pretrained on the grayscale CIFAR-10 dataset, and the training and fine-tuning were carried out on the real and simulated GPR data.

However, the existing research has some limits in automatic positioning study, and the performance of detection model can not meet the needs of obtaining the information on the location of underground objects accurately. On the basis of the above studies, this letter proposes a method based on deep learning for the automatic recognition and localization of underground pipelines using GPR.

II. Methodology

A. Recognition Model Based on Faster R-CNN

In the object detection algorithm, the operations of target classification and detection frame regression are completed on the basis of the feature map extracted by the backbone network, so the quality of the extracted image features will

Manuscript received 16 February 2022; revised 24 June 2022; accepted 8 August 2022. Date of publication 11 August 2022; date of current version 16 August 2022. This work was supported in part by the National Natural Science Foundation of China under Grant 51978630 and Grant 52108289, in part by the Postdoctoral Research Foundation of China under Grant 2020M672276 and Grant 2021T140620, in part by the Outstanding Young Talent Research Fund of Zhengzhou University under Grant 1621323001, in part by the Transportation Science and Technology Project of Henan Province under Grant 2018J7, and in part by the Key Scientific Research Project of Henan Province under Grant 21A560013. *(Corresponding authors: Hongyuan Fang; Niannian Wang.)*

Haobang Hu, Hongyuan Fang, Niannian Wang, Jianwei Lei, Duo Ma, and Jiaxiu Dong are with the School of Water Conservancy Engineering, Zhengzhou University, Zhengzhou 450001, China (e-mail: zzuhhb@gs.zzu.edu.cn; fanghongyuan@zzu.edu.cn; wnnian@zzu.edu.cn; 15333827823@163.com; md13255961153@gs.zzu.edu.cn; dongjiaxiu@gs.zzu.edu.cn).

Hai Liu is with the School of Civil Engineering, Guangzhou University, Guangzhou 510006, China (e-mail: hliu@gzhu.edu.cn).

Digital Object Identifier 10.1109/LGRS.2022.3198439

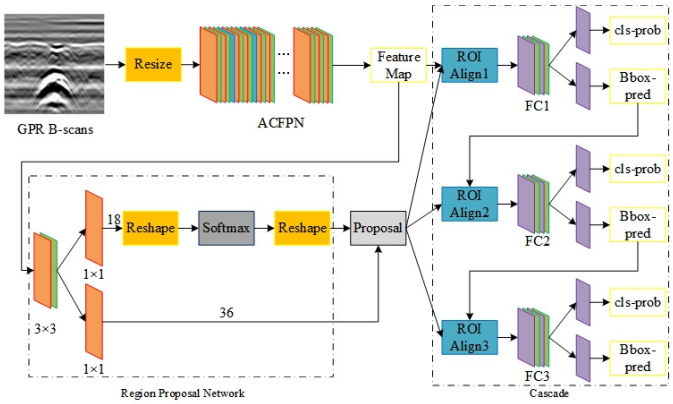

Fig. 1. Schematic of the proposed model.

directly affect the final detection accuracy. In addition, the regression detection frame Quality also has a great impact on detection accuracy. In summary, this study improves the model from the perspective of feature extraction and detection frame regression, and introduces the Attention-guided Context Feature Pyramid Network (ACFPN) and cascade detection structure to improve the accuracy of pipeline detection. As shown in Fig. 1, the network structure of the model can achieve the function of end-to-end training and testing [14].

1) ACFPN: Pipelines have the characteristics of various shapes and sizes, which are difficult to identify. After the image is input into the feature extraction network, several feature maps of different dimensions from low to high latitude will be generated. The feature maps of different dimensions contain different feature information. The FPN network can fuse the feature maps of different dimensions and use them for subsequent classification and classification. The regression operation enhances the feature expression ability of the CNN, but the Feature Pyramid Network (FPN) method still has some limits. When the resolution of the input image is too high, the receptive field cannot cover the entire image well, and effective semantic information cannot be obtained. Making full use of the environmental information around the target, the detection effect for large-scale targets is not effective enough. After multiple downsampling, the resolution of the feature map becomes smaller. Although the receptive field will become larger, the detailed features will be lost. Semantic information utilization is low.

In response to the above limits, this study introduces an ACFPN [15], which can fuse receptive field features from different sizes, increase the receptive field, and make full use of the contextual semantics of objects information for higher accuracy. The network structure of ACFPN is shown in Fig. 2. The network is mainly composed of two parts. One part is the Context Extraction Module (CEM), which is used to extract Context information from receptive fields of different sizes. The other part is the Attention-guided Module (AM), which can significantly enhance the useful information around the object through the attention mechanism, and avoid a large number of complex contextual relationships that mislead the localization and recognition task. The AM module contains two sub-modules: the Context Attention Module (CxAM) and the Content Attention Module (CnAM). The CxAM is used to capture the semantic information between the sub-regions, and

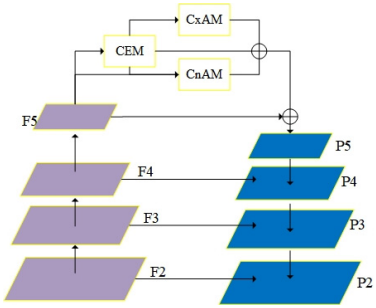

Fig. 2. Structure of ACFPN.

the CnAM is used to locate the precise location information. Fusion of the features refined by the two modules with the input features can obtain a more comprehensive feature representation, thereby obtaining better detection results.

2) Cascade Detection Structure: Intersection over Union (IOU) is an indicator used to evaluate the proximity of the predicted frame to the target frame. The larger the IOU value, the closer the predicted frame is to the real frame. The minimum value of IOU is 0, which means that the predicted frame and the real frame do not intersect. And the maximum value that can be taken is 1, which means that the predicted frame and the real frame coincide. In the classification and regression stage of the Faster R-CNN, an IOU threshold will be set to define positive and negative samples. If the threshold is set too small, the positive samples will also contain a lot of background information, which is likely to cause false detection. If the threshold is set too large, it will cause positive and negative samples. When the number of samples is too small, when the IOU value of the input proposal box and the real box is close to the set threshold, the detection effect of the model is the best [16]. In order to further improve the detection rate and accuracy of the detection frame and avoid the missed detection and the deviation of the detection frame from the real frame, the Cascade detection structure is introduced based on the Faster R-CNN and ACFPN. The Cascade structure is characterized by cascading multiple detectors with different IOU thresholds, using the output of the former detector as the input of the latter detector, and the IOU gradually rises. The IOU at the output of each level structure is generally larger than the IOU at the input, so the detection effect will gradually become better.

B. Localization Model Based on Tesseract OCR

This work employs the Tesseract OCR [17] for character recognition. The function realization process of the Tesseract OCR chiefly includes two parts: Analysis of GPR B-scans Layout, Segmentation and Recognition of Characters. The coordinates of the GPR B-scan image in the radar display interface are all black and white printed characters with the same format, and the background is pure white. All the characters are arranged horizontally, and the distance between each character block is large. Therefore, no specific language library is established for Arabic numerals. Operations such as noise processing and angle rotation correction are also unnecessary, which improves the efficiency and accuracy of character recognition and can meet the needs of pipeline positioning.

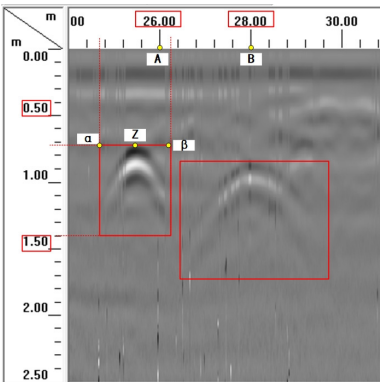

Fig. 3. Schematic of the test results. k is the coefficient correlating the objective real distance corresponding to the pixel with the difference in the pixel coordinates.

After the radar image data are imported into the algorithm processing module, the detection and recognition network model is utilized to classify and locate the pipeline features in the radar image. The positioning model attempts to determine the depth of burial (y) and the horizontal position (x), i.e., the horizontal distance from the pipeline to the starting point of the GPR line. In the detection and recognition network model, as shown in Fig. 3, the coordinate values of the left vertex (α) and the right vertex (β) of the rectangle can be output as $(p_{x\alpha}, p_{y\alpha})$ and $(p_{x\beta}, p_{y\beta})$ respectively. It is assumed here that the rectangular box happens to surround the pipeline feature, so the position of the midpoint of the left and right vertices (Z) can correspond to the condition of the buried pipeline. In this letter, RADAN7 software is utilized to process the radar image data. After the resolution of the display window is fixed, the coordinate point closest to α (A) and the coordinate point closest to β (B) on the transverse axis of the software display interface are obtained. (x_z, y_z) is expressed as follows:

$$\begin{cases} k = \dfrac{x_B - x_A}{p_{x_B} - p_{x_A}} \\ x_z = x_A + (p_{x_z} - p_{x_A}) \cdot k \\ y_z = (p_{y_z} - p_{y_o}) \cdot k \end{cases} \tag{1}$$

where (p_{x_i}, p_{y_i}) and (x_i, y_i) are the pixel coordinates and real coordinates of point i respectively. Point O is the origin of the orthogonal axis of the software interface; it corresponds to the ground, where the depth of burial is zero, and to the starting point of the GPR line, where the horizontal position is zero.

III. Datasets

Due to the existence of various noise sources around the underground pipelines and the complex underground environment, the quantity and quality of the GPR B-scans collected in the actual project are insufficient and poor and thus cannot satisfy the requirements of the model training. Therefore, a composite dataset of the GPR B-scans of underground pipelines (GBUP) including the real images, the case experimental images, and the images simulated by the finite difference time domain (FDTD) method [18] is established. The real images are collected with GSSI SIR-4000, using 200 MHz, 400 MHz, a total of two kinds of ground coupling antennas, and detection depth up to 1–5 m. The experiments are based on the measured GPR B-scans of the underground pipelines in Zhengzhou, China. Relative dielectric constant (epsr) is determined to be 10–12 according to local underground soil condition and B-scan quality in practical application. As shown in Fig. 4, there are 300 original images, and 2400 images are obtained after the DA. The study on the real-time positioning of the GPR B-scans is oriented to the detection of the practical use scenes. The real radar images account for 83.3% of the total images in the dataset, and the rest of the images are used to enrich the dataset and improve the generalization ability of the model.

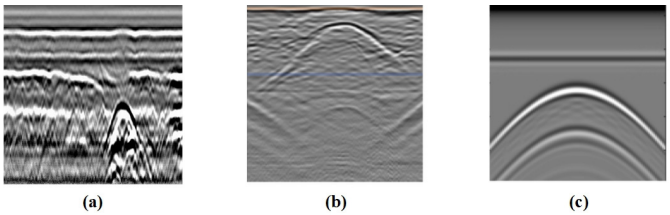

Fig. 4. Examples of the GPR B-scans. (a) Real radar images, (b) radar images of the case experiment, and (c) radar images simulated by the FDTD method.

The processing of B-scans includes: calibrating the surface layer so that the scale of the map is proportional to the depth of the ground; The dc offset of each signal is removed so that the normal value of the signal corresponds to zero; Bandpass filtering to remove possible high- and low-frequency noise; Background removal is used to remove the highlighting reflection caused by normal structural stratification and highlight abnormal signals; Gain is used to increase the magnitude of the deep reflected signal to the same extent as the shallow signal. The processing method is consistent within the same area. Then the open source labeling software was used to complete the data annotation. These operations are done manually.

IV. Experiments and Results

A. Comparison Experiment

The deep learning platform is convolutional architecture for fast feature embedding (Caffe). The modeling is conducted using a system equipped with Intel[1] Xeon(R) Gold 6138 CPU @ 2.00 GHz, 128.0 GB of memory, and NVIDIA Quadro P2000 graphics card. The GPR image test platform is built based on Ubuntu 16.04 operating system, Python programming language version 3.8. The detection experiments including randomly generating datasets, increasing or decreasing the capacity of the datasets, and increasing or decreasing the number of iterations are carried out. Pascal VOC-2007 is considered as the dataset of the pretraining model, and then the real and simulated datasets are used to fine-tune the parameters of the model based on the pretraining model. The loss value is output at intervals of 20 iterations at the terminal during the training process. To determine whether the model is optimal, its output at every 2500 iterations is used to detect the testing set. The better scheme which not only can ensure a high recognition accuracy but also can prevent the model training time from becoming too long is composed of the proposed model, 20 800 iterations, with an AP of 0.9256. To compare the performance of the proposed method with that of other

[1]Registered trademark.

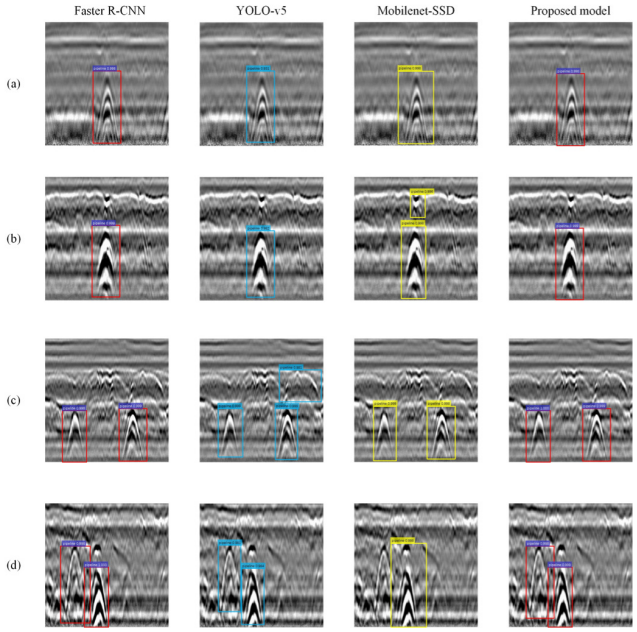

Fig. 5. Recognition effects of the different models. (a) Single metallic pipeline, (b) single nonmetallic pipeline, (c) multiple nonoverlapping features, and (d) multiple overlapping features.

methods, Faster R-CNN, YOLO-v5, and Mobilenet-SSD are used to detect the same radar image dataset.

Fig. 5 compares the test results of the four methods. Fig. 5(a) and (b) shows the detection effects of a single metallic pipeline and a single nonmetallic pipeline respectively. The area identified by YOLO-v5 and Mobilenet-SSD is larger; hence, it cannot effectively identify the feature boundary, and there are multiple inspection cases. Fig. 5(c) presents two nonoverlapping features. Fig. 5(d) depicts two overlapping features, and YOLO-v5 has significant error detection and multiple detection. Mobilenet-SSD miss the left pipeline feature; Faster R-CNN can identify two overlapping pipeline features and has neither error detection nor multiple inspection, but its recognition area is inaccurate. The model proposed can accurately identify the conditions of the four types of the buried pipelines, and the generated prediction box is more consistent with the location and size of the features. The prediction box surrounds the characteristic area of the underground pipeline, and the upper edge of the box coincides with the vertex position of the hyperbolic feature. The model developed herein is robust for detecting the pipeline characteristics under different burying conditions, including a complex cross-overlap. The model is also accurate, reliable, and can be well applied to the practical detection.

To further evaluate the proposed detection model, we compared Accuracy, Precision, Recall, F1-score, AP, and Fps, which are shown in Table I. The Accuracy, Precision, Recall, F1-score, and AP of the detection results of the proposed network on the image to be tested are higher than those of the other three models. Compared with the original Faster R-CNN, the performance of our proposed model is also slightly improved, but the detection speed is reduced. The accuracy of YOLO-V5 and Mobilenet-SSD is not as good as the improved scheme in this letter, but the detection speed FPS value far exceeds that of the two-stage target detection network.

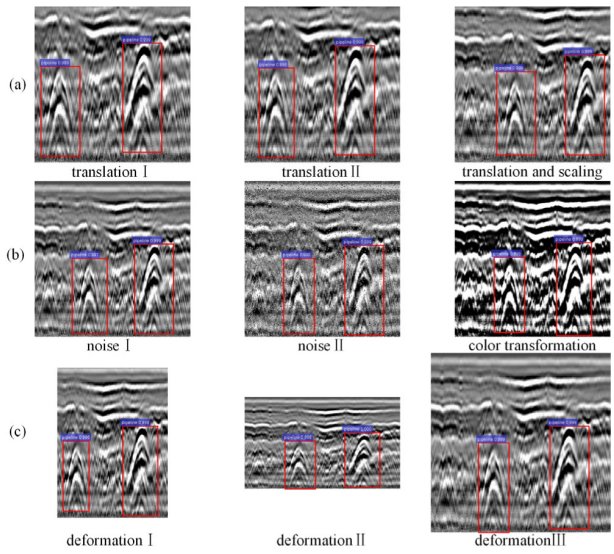

Fig. 6. Analysis of the robustness of the proposed model. (a) Translating and scaling, (b) noise and color transformation, and (c) deformation.

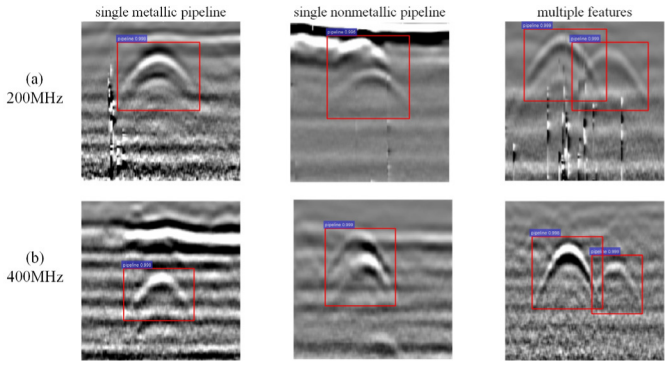

Fig. 7. Analysis of the robustness of different frequencies of transmitting antennas. (a) 200MHz and (b) 400MHz.

B. Robustness Testing

The robustness is verified from the aspects of translation, scaling, noise, color transformation, and deformation. Fig. 6(a), translating and scaling the image, indicates that the position of the pipeline features in the image and the size ratio relative to the whole image do not affect the recognition performance of the model. Fig. 6(b) adds Gaussian noise and color transformation, which can simulate the influence of different underground environments and noise. Fig. 6(c) depicts the proportional deformation of the image by stretching the pipeline features horizontally or vertically and confirms that the developed model can still detect the pipeline features successfully and accurately. Fig. 7 shows the test results of the proposed model on the corresponding data at 200 and 400 MHz. The results prove that the model has strong robustness and stability. It not only can recognize the pipeline characteristics of deformation or translation but also can adapt to a certain noise environment. Hence, it can be helpful in recognizing pipeline targets in images with more interference and blur which are difficult for human to distinguish.

C. Pipeline Locating

The depth of burial and horizontal position of the underground pipeline can be output with centimeter-level accuracy.

TABLE I
TEST RESULTS FOR DIFFERENT MODELS

Models	ACC	Precision	Recall	F1-score	AP	Fps
Faster R-CNN	0.897	0.918	0.843	0.879	0.9213	1.82
YOLO-v5	0.786	0.752	0.692	0.721	0.8317	23.54
Mobilenet-SSD	0.761	0.746	0.673	0.708	0.7924	24.32
Proposed model	0.921	0.963	0.897	0.929	0.9256	1.63

TABLE II
SOME OF THE RESULTS OF PIPELINE LOCATION INFORMATION

Scenario	x				y			
	x (cm)	x' (cm)	Absolute error (cm)	Error ratio	y (cm)	y' (cm)	Absolute error (cm)	Error ratio
(a)	92	81	11	12.0%	153	162	9	5.8%
(b)	156	165	9	5.7%	174	164	10	5.7%
(c)	364	354	10	2.7%	203	211	8	3.9%
(d)	410	399	11	2.6%	286	278	8	2.7%

In the detection of urban municipal drainage pipelines, by opening the manhole cover for manual measurement, 50 sets of the output results are compared with the known actual information on the buried pipeline. Some of the result are shown in Table II. The absolute error in the depth of burial and the horizontal position is lower than 11 cm, and the average error ratio is smaller than 12%.

V. CONCLUSION

This letter proposes an accurate and efficient model based on Faster R-CNN to tackle the challenge of recognition buried pipelines using GPR B-scan data. The feature extraction is optimized by ACFPN, and the cascade structure is used to improve the detection frame regression accuracy. Then, a positioning model is developed based on the recognition result using the Tesseract OCR to obtain the burial and horizontal position of buried pipelines. The results demonstrate that the average precision (AP) of the recognition can reach 0.9256.

Furthermore, a combination of multigroup model training and application testing is designed to make the model have good recognition and noise robustness. The method is simple and easy, and it is less disturbed by the size and rotation angle of radar images. The comparison of the obtained results with Faster R-CNN, YOLO-v5, and Mobilenet-SSD also demonstrates the superiority of the established recognition model.

The absolute error of burial and horizontal position prediction is lower than 11 cm, and the average error ratio is smaller than 12%. The method can greatly enhance the efficiency of underground pipelines detection based on GPR, and can provide database support for the pipeline census, and can contribute to the practical applications.

In addition, since there are a variety of underground buried targets such as void, loose and other roadbed defects, we will collect different types of GPR data to train the developed model so as to broaden the field of intelligent identification.

REFERENCES

[1] B. Yektakhah, J. Chiu, F. Alsallum, and K. Sarabandi, "Low-profile, low-frequency, UWB antenna for imaging of deeply buried targets," *IEEE Geosci. Remote Sens. Lett.*, vol. 17, no. 7, pp. 1168–1172, Jul. 2019.

[2] A. Simi, S. Bracciali, and G. Manacorda, "Hough transform based automatic pipe detection for array GPR: Algorithm development and on-site tests," in *Proc. IEEE Radar Conf.*, May 2008, pp. 1–6.

[3] S. W. Jaw and M. Hashim, "Urban underground pipelines mapping using ground penetrating radar," in *Proc. IOP Conf. Earth Environ. Sci.*, vol. 18, Feb. 2014, Art. no. 012167.

[4] G. Jiang, X. Zhou, J. Li, and H. Chen, "A cable-mapping algorithm based on ground-penetrating radar," *IEEE Geosci. Remote Sens. Lett.*, vol. 16, no. 10, pp. 1630–1634, Oct. 2019.

[5] J. Li, C. Liu, Z. Zeng, and L. Chen, "GPR signal denoising and target extraction with the CEEMD method," *IEEE Geosci. Remote Sens. Lett.*, vol. 12, no. 8, pp. 1615–1619, Aug. 2015.

[6] E. Pasolli, F. Melgani, and M. Donelli, "Gaussian process approach to buried object size estimation in GPR images," *IEEE Geosci. Remote Sens. Lett.*, vol. 7, no. 1, pp. 141–145, Jan. 2010.

[7] I. Giannakis, A. Giannopoulos, and C. Warren, "A machine learning scheme for estimating the diameter of reinforcing bars using ground penetrating radar," *IEEE Geosci. Remote Sens. Lett.*, vol. 18, no. 3, pp. 461–465, Mar. 2020.

[8] L. E. Besaw and P. J. Stimac, "Deep convolutional neural networks for classifying GPR B-scans," *Proc. SPIE*, vol. 9454, pp. 385–394, May 2015.

[9] A. Krizhevsky, I. Sutskever, and G. E. Hinton, "ImageNet classification with deep convolutional neural networks," in *Proc. Adv. Neural Inf. Process. Syst.*, vol. 25, 2012, pp. 1097–1105.

[10] K. Dinh, N. Gucunski, and T. H. Duong, "An algorithm for automatic localization and detection of rebars from GPR data of concrete bridge decks," *Automat. Construct.*, vol. 89, pp. 292–298, May 2018.

[11] D. Reichman, L. M. Collins, and J. M. Malof, "Some good practices for applying convolutional neural networks to buried threat detection in ground penetrating radar," in *Proc. 9th Int. Workshop Adv. Ground Penetrating Radar (IWAGPR)*, Jun. 2017, pp. 1–5.

[12] M.-T. Pham and S. Lefevre, "Buried object detection from B-scan ground penetrating radar data using faster-RCNN," in *Proc. IEEE Int. Geosci. Remote Sens. Symp. (IGARSS)*, Jul. 2018, pp. 6804–6807.

[13] C. Warren, A. Giannopoulos, and I. Giannakis, "GprMax: Open source software to simulate electromagnetic wave propagation for ground penetrating radar," *Comput. Phys. Commun.*, vol. 209, pp. 163–170, Dec. 2016.

[14] S. Ren, K. He, R. Girshick, and J. Sun, "Faster R-CNN: Towards real-time object detection with region proposal networks," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 39, no. 6, pp. 1137–1149, Jun. 2017.

[15] J. Cao, Q. Chen, J. Guo, and R. Shi, "Attention-guided context feature pyramid network for object detection," 2020, *arXiv:2005.11475*.

[16] M. Ahmad, I. Ahmed, and G. Jeon, "An IoT-enabled real-time overhead view person detection system based on cascade-RCNN and transfer learning," *J. Real-Time Image Process.*, vol. 18, no. 4, pp. 1129–1139, Aug. 2021.

[17] R. Smith, "An overview of the tesseract OCR engine," in *Proc. 9th Int. Conf. Document Anal. Recognit. (ICDAR)*, Sep. 2007, pp. 629–633.

[18] D.-S. Feng and Q.-W. Dai, "GPR numerical simulation of full wave field based on UPML boundary condition of ADI-FDTD," *NDT & E Int.*, vol. 44, no. 6, pp. 495–504, Oct. 2011.