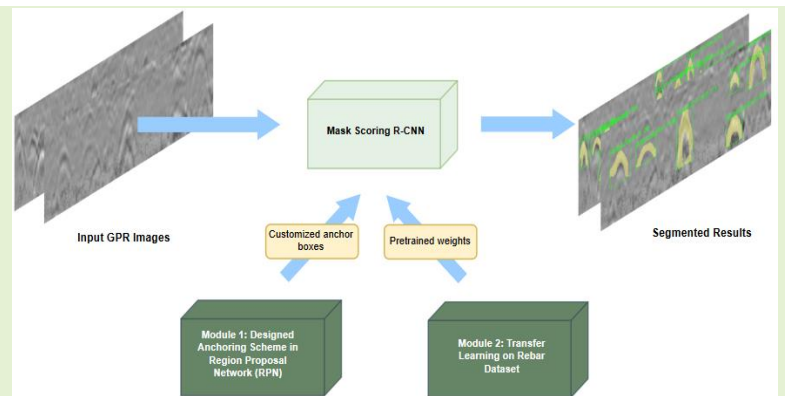


Deep Learning Based Subsurface Target Detection from GPR Scans

Feifei Hou, Wentai Lei, Shuai Li, and Jingchun Xi

Abstract—Ground penetrating radar (GPR) has been widely used as a non-destructive technique to detect subsurface objects. Manual process and interpretation of GPR data is tedious and time-consuming. To address this challenge, an automatic method based on a deep instance segmentation framework is developed to detect and segment object signatures from GPR scans. The proposed method develops the Mask Scoring R-CNN (MS R-CNN) architecture by introducing a novel anchoring scheme. By analyzing the characteristics of the hyperbolic signatures of subsurface objects in GPR scans, a set of anchor shape ratios are optimized and selected to substitute the predefined and fixed aspect ratios in the MS R-CNN framework to improve the signature detection performance. In addition, transfer learning technique is adopted to obtain a pre-trained model to address the challenge of insufficient GPR dataset for model training. The detected and segmented signatures can then be further processed for target localization and characterization. GPR data of tree roots were collected in the field to validate the proposed methods. Despite the noisy background and varying signatures in the GPR scans, the proposed method demonstrated promising results in object detection and segmentation. Computational results show that the improved MS R-CNN outperforms the other state-of-the-art methods.

Index Terms—Ground penetrating radar (GPR), deep learning (DL), instance segmentation, root detection, Mask Scoring R-CNN (MS R-CNN), anchor box.



I. INTRODUCTION

Ground penetrating radar (GPR) have been adopted for subsurface mapping as it could provide rich information regarding the objects buried at varying depths in the subsurface [1]. GPR has been used in a variety of domains ranging from civil engineering [2-5], archaeology [6, 7] to agricultural activities [8-12]. In most of the applications such as non-destructive bridge assessment [13, 14], underground utility mapping [15, 16], and tree root detection [17, 18], detecting and identifying of objects from GPR scans is a vital step to retrieve meaningful information. However, manual processing GPR data is time-consuming and labor-intensive, and thus are not suitable for large amounts of GPR data acquired from the field.

Manuscript received XX; revised XX; accepted XX. Date of publication XX. This work was funded by the National Natural Science Foundation of China (nos. 61102139), the National Science Foundation of U.S (nos. 1850008), and the China Scholarship Council (CSC). The authors wish to thank Mr. Da Hu (from University of Tennessee, Knoxville, US) for helping collect field data and Mr. Mengdi Xu (from Central South University, China) for supporting the interpretation of the results. (Corresponding author: Shuai Li)

Feifei Hou, Wentai Lei, and Jingchun Xi Authors are with School of Computer Science and Engineering, Central South University, Changsha, China (e-mail: houfeifei@csu.edu.cn; leiwentai@csu.edu.cn; xijingchun@csu.edu.cn). Shuai Li and Feifei Hou Authors are with Department of Civil and Environmental Engineering, University of Tennessee, Knoxville, USA (e-mail: sli48@utk.edu).

Therefore, an automatic method is needed for accurate detection of subsurface objects from GPR scans.

The existing studies have demonstrated the feasibility of using GPR to scan the subsurface and its promise for extracting useful information from GPR profiles to locate objects in subsurface. Li et al. [17] used randomized Hough transform method to detect tree roots in GPR scans. The performance was evaluated in controlled and in situ experiments, where root datasets were collected by GPR with different center frequencies. The method presented in [19] first identifies hyperbola regions in GPR scans, and then uses the generalized Hough transform to locate the hyperbolic signatures of the buried objects. Studies [20, 21] used a genetic algorithm (GA)-based technique to identify the linear and hyperbolic features resulted from subsurface objects in binary GPR images. Based on the extracted features, Harkat et al. [21] further used a classifier based on neural network radial basis function (RBF) to classify windows of GPR radargrams into two classes: regions with and without objects. The success of technique presented in [22] demonstrated the potential of using histogram of oriented gradient (HOG) feature extraction and support vector machine (SVM) to detect object signatures in GPR scans. However, most input features need to be recognized by practical experts, and the classification result depends on the quality of these features, which are limited by the amount of GPR data processed.

In recent years, deep learning (DL) techniques have been exploited to automatically detect object signatures in GPR scans [23-25]. DL methods can learn the feature representation of buried objects from GPR data. In the method developed by Kien Dinh et al. in [26], first image processing operations (zero-time correction, migration, filtering, and threshold segmentation) are used to find the vertices of each hyperbola. The fixed-size areas around vertices are then extracted and divided into positive or negative sample, which are used as the input for convolutional neural networks (CNN) to train the model. Finally, the trained model is used to test the candidate dataset. The work in [27] proposed a hybrid deep CNN to detect and locate moisture damage of bridges from GPR data. This hybrid model consists of the ResNet50 [28] network for feature extraction and YOLO v2 [29] network for target recognition. The typical object detection framework, Faster R-CNN [30], is used in [24] to classify buried linear objects and narrow down candidate hyperbola region. The hyperbolic clusters are segmented from complex background using a double cluster seeking estimate (DCSE) algorithm. In addition, a novel column-based transverse filter points (CTFP) method is used to automatically extract data points from hyperbolic regions. However, complete information could not be extracted from the detected hyperbolic signatures in that study. To address issue, an enhanced DL framework is proposed in [31], which integrates Mask R-CNN [32] and a novel loss function computation to simultaneously detect and segment hyperbolic signatures. Instead of using clustering-based methods to obtain segmented results as in [24, 33, 34], both the box and mask patches of hyperbolic signatures can be obtained.

Although the application of DL techniques in GPR data processing has achieved partial success, the performance of DL models still remains a concern in real-world applications. First, the formation of hyperbolic signatures in GPR depends on a variety of conditions, including the size of the buried object, buried depth, antenna frequency, and soil condition. Due to the complex subsurface conditions and the radar wave reflections, the object signatures in GPR images could also be incomplete and distorted. This imposes difficulties on developing a DL

model that could achieve desired performance for specific GPR application. Second, DL model can be adopted for most image-related tasks. However, radargram is different from conventional images. The internal architectures of DL models have not been well linked with the attributes and characteristics of radargram and the object signatures. Thus, direct application of such DL models for GPR data processing may omit useful information and lead to redundant processing.

We proposed a new method to address the limitations, as shown in Fig. 1. First, a DL-based framework was developed to automate the detection and segmentation of object signatures in GPR scans. The innovation is twofold. The Mask Scoring R-CNN (MS R-CNN) [35] predefines a set of fixed anchor shapes (e.g. aspect ratios, scale) to generate proposals. However, using these generated proposals is not suitable for detecting various object signatures with different sizes and shapes in GPR scans. Therefore, we first analyzed the hyperbolic signatures and develop a new customizable anchoring scheme to enhance the proposal generation to improve detection performance for GPR scans. In addition, DL-based methods need a large number of labeled datasets to train the model. Unlike conventional image processing tasks, the labeled GPR datasets are very scarce. We adopt the transfer learning method to solve the problem of inadequate training dataset and improve the model robustness. The new anchoring scheme and the transfer learning method are integrated with the MS R-CNN to process GPR scans to detect and segment object signatures. Second, the segmented signatures are extracted from cluttered background, and then hyperbolic fitting is performed to find the peak that approximately indicate the object position.

The rest of the paper is organized as follows. Section II designs a novel anchoring scheme and incorporates it into the MS R-CNN framework. Section III elaborates the customized anchors for subsurface targets and transfer learning technique, and introduces the automatic target recognition scheme based on the improved MS R-CNN framework. Section IV presents experimental results, followed by the conclusion and future research in Section V.

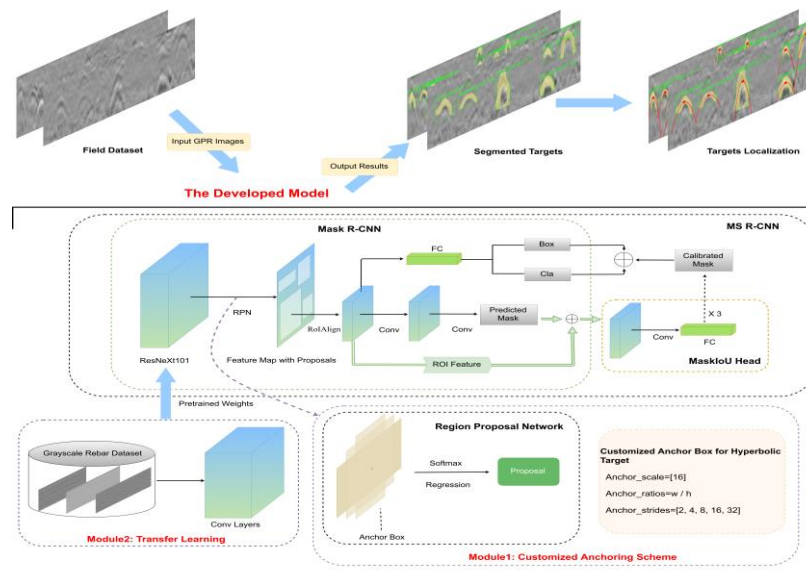


Fig. 1. The proposed GPR object signature detection and segmentation framework.

II. IMPROVED MS R-CNN FRAMEWORK

A new anchoring scheme and a transfer learning technique are developed and integrated into the MS R-CNN to automatically detect and segment hyperbolic signatures of buried objects from GPR scans. Table 1 lists the abbreviations.

TABLE I
SUMMARY OF ABBREVIATIONS

Abbreviation	Description
AP	Average Precision
AR	Average Recall
bbox	Bounding Box
CNN	Convolutional Neural Networks
DL	Deep Learning
FPN	Feature Pyramid Network
FCs	Fully Connected Layers
GPR	Ground Penetrating Radar
GT	Ground Truth
IoU	Intersection over Union
NMS	Non-maximum Suppression
RPN	Region Proposal Network
RoI	Region of Interest
RoIAlign	Region of Interest Align
R-CNN	Region-based CNN
Soft-NMS	Soft Non-maximum Suppression

A. Overview of MS R-CNN

The instance segmentation model can simultaneously classify, detect, and segment objects in the input image. For most instance segmentation tasks, the quality of segmented mask is determined by the classification confidence of the detection branch. However, the mask quality does not necessarily match with the classification confidence. MS R-CNN is a typical instance segmentation framework to address this issue [35]. It integrates an additional network block ‘MaskIoU head’ into Mask R-CNN to explore the mask scoring. Its network architecture consists of three stages as denoted in Fig. 1. First, a region proposal network (RPN) is leveraged to generate a set of candidate proposals. The second stage is to use region of interest align (RoIAlign) to extract features from each proposal and feed them into two standard components: region-based convolutional neural networks (R-CNN) head and Mask head. The former generates classification result and bounding box (bbox) regression result, while the latter performs mask prediction. The top-k (for example, top100) score bboxes output by R-CNN head are processed by soft non-maximum suppression (Soft-NMS), and then sent to Mask head to get masks. The third stage concatenates the obtained mask and RoI feature map into MaskIoU head to get MaskIoU value.

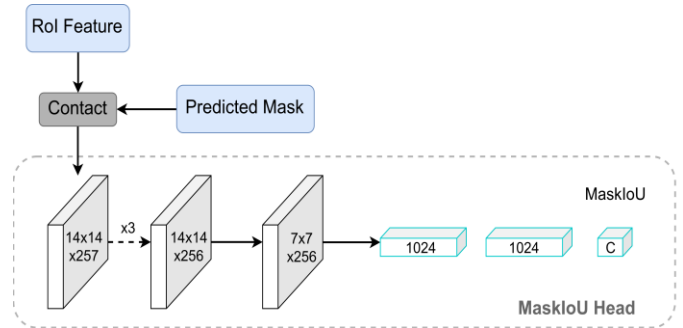


Fig. 2. MaskIoU head and its two inputs: predicted mask and RoI feature.

Fig. 2 shows the network architecture of the MaskIoU head branch. It includes four convolutional layers and three fully connected layers (FCs). The kernel size is set to 3 for all the convolutional layers, and stride is set to 2 for downsampling in the last convolutional layer. The final FC outputs C classes of MaskIoU. The MaskIoU head takes both predicted mask (generated via Mask head) and IoU features (yielded via RoIAlign) as inputs to re-calibrate mask. There are two tasks: classifying each mask into correct categories, and regressing the MaskIoU of each proposal for category correction. To perform these two tasks, the calibrated mask score is computed by multiplying classification confidence S_{cla} and IoU regression value S_{iou} , as in Equation (1).

$$S_{mask} = S_{cla} \cdot S_{iou} \quad (1)$$

S_{cla} is used to classify each candidate proposal into a certain category at the R-CNN stage, and S_{iou} is used to regress IoU value between the predicted mask and ground truth (GT) mask in MaskIoU head module. MaskIoU head is concatenated in the R-CNN head and the mask head of Mask R-CNN.

B. Designed Anchor Scheme

1) *Overview of Anchor*: Most state-of-the-art object detectors provide an anchor scheme that contains a large number of dense anchors. Anchors are used to generate proposals, extract the corresponding region features and regress the location of bboxes. Because the candidate targets may be unpredictably distributed in an input image with various shapes, a set of anchor boxes need to be generated based on the center point of every sliding window in feature map and is determined by predefined anchor scales and aspect ratios. The feature map, as the output of CNN, is used to roughly locate targets. The anchor scales and anchor aspect ratios are used to respectively represent the size and shape of target.

Modern object detectors are roughly divided into two-stage and single-stage pipelines. The two-stage approach is mainstream in object detection task. Generating anchors via a sliding window located in feature maps has been widely accepted in anchor-based object detectors. For instance, Faster R-CNN [30] innovatively proposed the RPN model to produce proposals from candidate anchors, and then these generated proposals are classified and regressed. These anchors are obtained by scanning feature map using the sliding window.

The Faster R-CNN defines three scales [8,16,32] and three ratios [0.5,1.0,2.0], combining 9 anchor boxes in different shapes. The Mask R-CNN and the MS R-CNN adopt the same design as Faster R-CNN. They define 15 anchor boxes with five scales [32,64,128,256,512] and three ratios [0.5,1.0,2.0]. In contrast to two-stage frameworks, the single-stage approach just classifies object to a certain category and predict bbox with confidence in one evaluation, and skips the stage of generating object proposal. The single-stage approach still requires the use of the sliding window to produce anchor boxes. For instance, the SSD [36] sets five different aspect ratios $[1,2,3,\frac{1}{2},\frac{1}{3}]$ to

generate a series of anchors. Facing the challenge of balancing the proportion of positive and negative samples, the RetinaNet [37] approach introduces focal loss into backbone to lower the weight of negative samples in training process to address the above challenge. Specifically, YOLOv2 [29] and YOLOv3 [38] utilize K-means [39] method to learn and yield different anchors from training set, rather than the combination of scales and aspect ratios.

2) Customized Aspect Ratios based on Hyperbolic Signatures: The essence of object detection task in GPR B-scans is to identify hyperbolic signature. The formation of hyperbolic signatures in GPR scans depends on the subsurface object (e.g., object size), underground environment (e.g., soil dielectric constant), and antenna frequency. Therefore, the object signatures are of various shapes and sizes. Applying the fixed anchoring scheme on different types of buried objects or on different handling of problems will generate many redundant proposals, which compromise the detection and segmentation performance. Hence, a new anchoring scheme is needed.

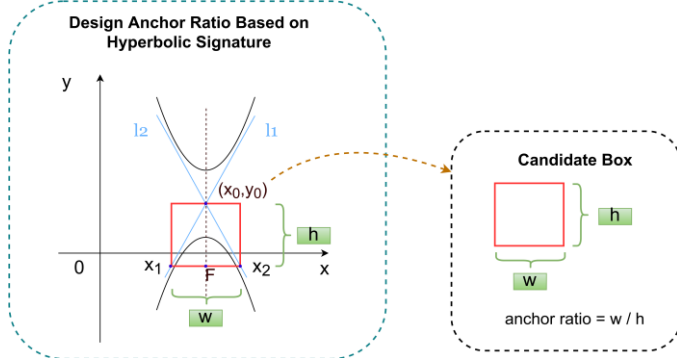


Fig. 3. Sketch of anchor ratio computation based on hyperbolic signature.

In this study, we customize the specific anchors based on the property of hyperbolic curves of buried objects. First, a concept of ‘candidate box’ is given, which is a rectangle defined by the location of focus and symmetric center (x_0, y_0) of a hyperbola, as well as asymptotes l_1, l_2 . The aspect ratio of the candidate box is computed to approximately represent the shape of hyperbola. The computed aspect ratio will replace the previous fixed aspect ratios in the MS R-CNN. Fig. 3 illustrates a hyperbola and the candidate box for computation of the aspect ratios. The general equation of hyperbola is denoted in Equation (2) [40]:

$$\frac{(y-y_0)^2}{a^2} - \frac{(x-x_0)^2}{b^2} = 1, (a > 0, b > 0) \quad (2)$$

Where (x, y) represents the arbitrary position in hyperbola, (x_0, y_0) is the symmetric center of hyperbolic curve. a and b denote solid semi-axis and imaginary semi-axis respectively. Equation (3) indicates the corresponding expansion of Equation (2). Also, Equation (4) [40] gives its asymptote.

$$-a^2 \cdot x^2 + b^2 \cdot y^2 + 2a^2x_0 \cdot x - 2b^2y_0 \cdot y + b^2y_0^2 - a^2x_0^2 - a^2b^2 = 0 \quad (3)$$

$$y - y_0 = \pm \frac{a}{b}(x - x_0) \quad (4)$$

We assure the accurate location of F is $(x_0, y_0 - c)$, where $c = \sqrt{a^2 + b^2}$. Taking the ordinate of F , i.e., $y = y_0 - c$, as input of Equation (4), we can obtain Equation (5). The line where F is located and paralleling to X-axis intersects with two asymptotes l_1 and l_2 , and then two intersection points of x_1 and x_2 are obtained in Equation (6).

$$\begin{cases} l_1: & y_0 - c = \frac{a}{b} \cdot x_1 + \frac{a}{b} \cdot y_0 + x_0 \\ l_2: & y_0 - c = -\frac{a}{b} \cdot x_2 - \frac{a}{b} \cdot y_0 + x_0 \end{cases} \quad (5)$$

$$\begin{cases} x_1 = x_0 - \frac{b}{a} \cdot c \\ x_2 = x_0 + \frac{b}{a} \cdot c \end{cases} \quad (6)$$

The two vital parameters: width w and height h that belong to candidate box are computed and given in Equation (7), respectively. The ratio of w to h is computed as candidate anchor ratio for buried object, shown in Equation (8). Thus, the properties of hyperbolic signature, a and b , can be used to directly estimate the aspect ratios of anchors belonging to different targets.

$$w = |x_2 - x_1| = \frac{2bc}{a}, \quad h = c \quad (7)$$

$$ratio = \frac{w}{h} = \frac{2b}{a} \quad (8)$$

C. Transfer Learning based on ResNeXt101 in MS R-CNN

Training the DL model requires a large amount of labeled positive/negative samples. The limited GPR datasets are not adequate to support the model training procedure. In this section, to strengthen the imagery recognition performance and lower the demand for a large number of labeled samples, a transfer learning technique is adopted in the improved MS R-CNN model.

During pre-training phase, we first give two concepts: ‘source set’ and ‘target set’. The source set is selected to pre-processing model. The target set is used for experimental aim and presentation. First, a DL model is pre-trained based on the source set and the model information is obtained with updated weights and network parameters. Second, this pre-trained model is used to promote the training procedure on the target problem. The effectiveness of pre-training relies largely on the degree of similarity between source set and target set. Due to the complexity and uniqueness of GPR characteristics, it is not easy to search for a suitable and similar source set for GPR target set. We recognize this challenge and investigate its acceptable source set for model pre-training.

Inspired by [41], two criteria are used to determine the source set for pre-training the MS R-CNN framework. First, the dataset needs to be large enough to train the framework. Second, the source set should be as similar as possible to our target dataset in terms of the problem and the data properties. Hence, a GPR dataset was selected from [31] as the source set of this work, which was used to identify rebars in grey GPR scans for bridge deck assessment. This source set was collected on a concrete bridge deck that consists of 95 grey images with a size of 512×1676 and contains thousands of annotated object instances. These instances belong to the category of hyperbolic signatures.

Fig. 4 illustrates the transfer learning process from three aspects: input, processing, and output. The ResNeXt101 [42] model is first trained using both ImageNet dataset and bridge rebar dataset. Then a set of pre-trained neuron parameters and weights are generated. These parameters are added to the training process to obtain the final model, which can predict the location of multiple boundary boxes and multiple categories of targets in real time.

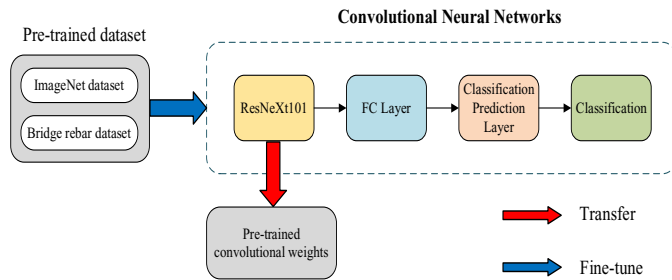


Fig. 4. Flowchart of the transfer learning based on ResNeXt101 network.

III. EXPERIMENTS AND IMPLEMENTATION

A. Field Data Collection

The field experiment was conducted in the UT Gardens ($35^{\circ}56'54.6''\text{N}$ $83^{\circ}55'52.1''\text{W}$), Knoxville, USA [Fig. 5]. By using the GSSI GPR SIR-4000 system with a center frequency of 2GHz, we collected radargrams of tree roots and will use it to test our developed method [Fig. 6]. The energy of a 2 GHz antenna can penetrate to a depth of about 0.5 m.



Fig. 5. Site aerial view and field site.

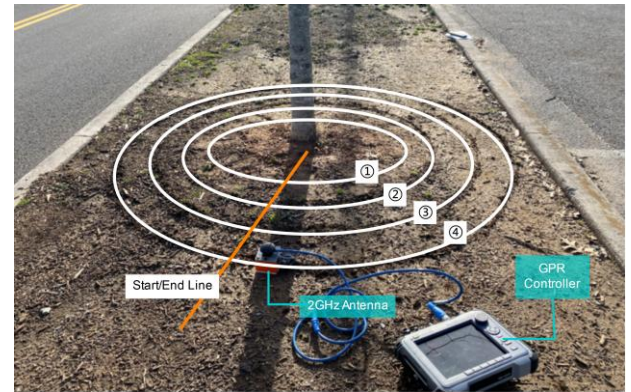


Fig. 6. In situ tree root experiment. Figure tabs the scan circle lines, start/end line, GPR control unit, and 2GHz antenna.

B. Experimental Setup

We used the PyTorch1.4.0 implementation of the MS R-CNN, which is based on the open-source object detection toolbox, namely mmdetection, released by Facebook research (https://github.com/zjhuang22/maskscoring_rcnn). All tasks are implemented using the computation environment with Nvidia GeForce GTX 1060 GPU. The ResNeXt-101 with feature pyramid network (FPN) is selected as the backbone network architecture for the MS R-CNN.

The source set has 95 GPR bridge scans containing thousands of annotated rebar instances. The target set includes 93 GPR root scans. To estimate the parameters and avoid overfitting, 10-fold cross validation is accepted by using 85% of the target data as training and 15% as testing at each fold. All input images are rescaled to $[800, 300]$ for both training and testing pipeline. We optimize the learning rate to 0.0025 and run 20000 iterations. In model testing configuration stage, we evaluate an appropriate threshold value of 0.5 for NMS in RPN stage and a score threshold of 0.8 in R-CNN stage. In RPN stage, the anchoring scheme is updated with $aspect_scale = [16]$, $aspect_ratios = [0.3492, 0.5406, 0.8528]$, $aspect_strides = [2, 4, 8, 16, 32]$.

C. Improved MS R-CNN Based Signature Detection

First, background removal and broad band filtering are applied to preprocess the raw GPR data. Then, the work of automatic object recognition is to: (1) customize anchors based on hyperbolic signature of root targets and integrate them into the instance segmentation model; (2) detect and segment hyperbolic signatures using the improved model; (3) identify targets based on curve fitting and peak localization.

1) *Customized Anchors for Targets*: Based on the observation of the hyperbolic signature of root targets, the target signature is relatively small compared to the entire GPR image, which leads to a small proportion of target in the image. This indicates a set of small-scale anchors needs to be designed for matching the root targets. We design a four-step computation method to obtain the customized aspect ratios instead of the previous fixed ones. The flowchart is shown in Fig. 7.

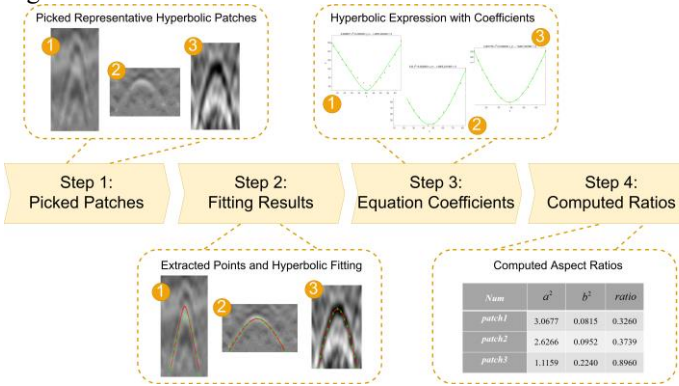


Fig. 7. Flowchart of the designed anchoring scheme for root objects. Step1: Example of picked representative patches (each patch contains a hyperbola); Step2: Extracted data points and the corresponding hyperbolic fitting; Step3: General hyperbola expression with coefficients; Step4: Parameter list and the computed aspect ratios of each patch.

As shown in Fig. 7, there are four stages to customize anchors for root objects. The first step is to pick several representative root samples from GPR images for subsequent aspect ratios computation. Two criteria are designed for selecting these samples: (1) the shape of each sample is distinct and inconsistent with other samples; (2) each sample contains only an obvious hyperbola. We select 7 different patch samples and manually select midpoints from each sample. These points are denoted as $P = [(x_1, y_1), (x_2, y_2), \dots, (x_i, y_i)]$, $15 \leq i \leq 25$, used for hyperbolic fitting. Based on this, the correlation coefficients (a and b) are obtained from the given hyperbola expansion and used to compute aspect ratio of the candidate box of each sample. Table 2 shows the details of each selected sample along with the corresponding hyperbolic coefficients and computed aspect ratios. According to the obtained 7 ratios, we roughly divide them into three groups: $R = [R_1, R_2, R_3]$, where $R_1 = [0.3260, 0.3476, 0.3739]$, $R_2 = [0.5773, 0.5039]$, $R_3 = [0.8960, 0.8096]$. Then, the average of each group is computed and we obtain three representative aspect ratios $[0.3492, 0.5406, 0.8528]$ and integrate them into the MS R-CNN framework as the updated anchors.

TABLE II

DETAILS OF THREE GROUPS INCLUDE SELECTED SAMPLES, HYPERBOLIC PARAMETERS, AND ASPECT RATIOS.

Group 1			Group 2		Group 3	
1	2	3	4	5	6	7
$a^2 = 3.0677$	$a^2 = 2.8775$	$a^2 = 2.6226$	$a^2 = 1.7320$	$a^2 = 1.9846$	$a^2 = 1.1159$	$a^2 = 1.2353$
$b^2 = 0.0815$	$b^2 = 0.0869$	$b^2 = 0.0952$	$b^2 = 0.1443$	$b^2 = 0.1260$	$b^2 = 0.2240$	$b^2 = 0.2024$
ratio = 0.3260	ratio = 0.3476	ratio = 0.3739	ratio = 0.5773	ratio = 0.5039	ratio = 0.8960	ratio = 0.8096

2) *Root Target Segmentation and Localization*: To obtain useful information, it is crucial to detect and segment targets from GPR scans and localize their peaks. First, the improved MS R-CNN model is used to detect and segment most hyperbola instances from complex background. After that, based on the previous work [31], we pick the segmented mask patches and extract data points from them for hyperbolic curve fitting and object peak localization. After that, the real location of underground coarse roots can be evaluated and analyzed.

IV. RESULTS AND DISCUSSION

A. Experiment Results

Compared with the controlled experiments, GPR measurement under field condition contains more and stronger noise and interference. The conventional field trials were conducted in dry sandy soil like in [17], where soil background is relatively homogeneous. However, our field data was collected under a more complex environment, which is a wet soil condition. Underground targets mainly are root targets and some disturbances, such as small stone, soil block or soft sand. To balance efficiency and accuracy, the detection task is focused on all underground targets. Since underground targets are often reflected as hyperbolic signatures in GPR profiles, underground target detection can be converted to the hyperbolic signature detection in GPR profile.

An example of recognition results for the coarse roots is given in Fig. 8. The pre-processed GPR B-scan is shown in Fig. 8(a). In addition to the several obvious hyperbolae distributed in middle of the figure, there are many subtle interferences distributed on the top, and the black blocks distributed on the bottom caused by the change of the medium. Fig. 8(b) shows the detection and segmentation results output by the enhanced MS R-CNN, including confidence, green bbox, and yellow mask region. Fig. 8 presents the separated mask patches on a black background, which provides target clusters information to simplify subsequent processing, such as data points extraction, curve fitting and peak locating [Fig. 8(d)]. It can be found that the improved MS R-CNN can pick most obvious and complete hyperbolae and is insensitive to small and dense hyperbolae. Fig. 9 provides more details about the intermediate procedures, including the acquisition of final localization results [see Fig. 8(d)] from each individual mask patch [see Fig.

8(c)]. To achieve the goal, operations are conducted to 1) extract all mask patches from Fig. 8(c) [Fig. 9(a)]; 2) threshold these patches into binary images [Fig. 9(b)]; 3) use the combination of opening and closing operations to process these binary images [Fig. 9(c)]; 4) eliminate non-target interferences and thus extract data points from target region [Fig. 9(d)]; 5) obtain hyperbolic fitting results with peaks, indicating the position of underground targets [Fig. 9(e)].

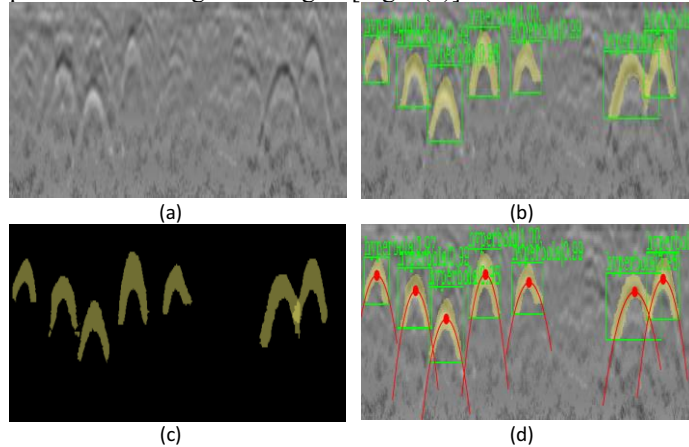


Fig. 8. First example of GPR root localization results processed by the proposed method. (a) Pre-processed root GPR scan. (b) Detected and segmented results. (c) Only picked mask clusters. (d) Hyperbolic fitting results with peaks (red point).

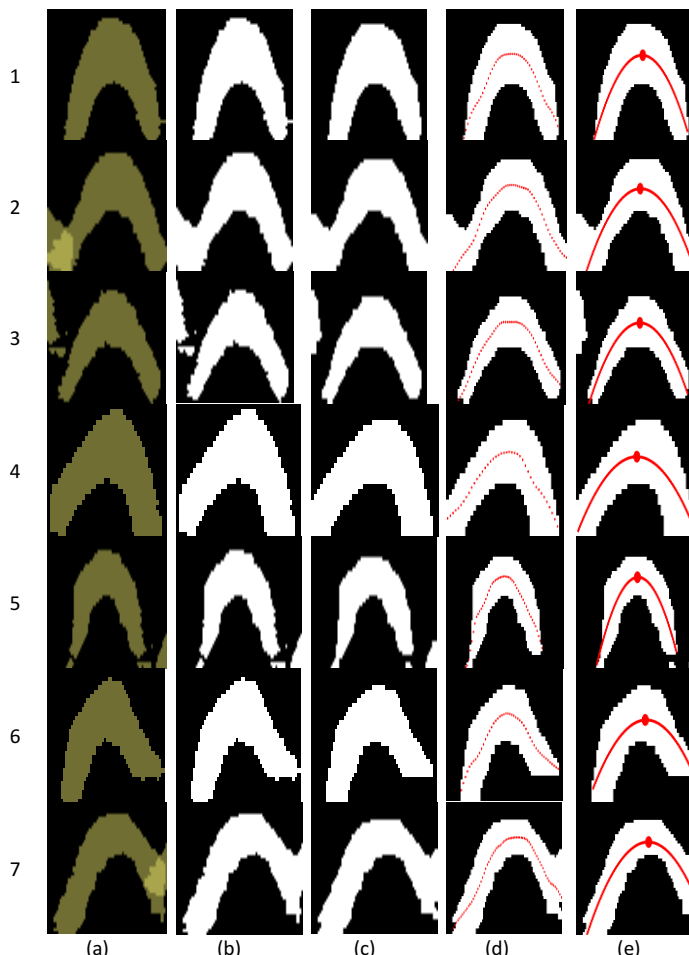


Fig. 9. The display of intermediate process. Group 1-7: (a) Segmented mask patches from Fig. 8(c); (b) Binary threshold results; (c)

Morphology operation results; (d) Extracted data points; (e) Fitting results.

The root signatures in the first scenery are clear and close to linear rebar or pipe signatures [31, 43], which is relatively easy to identify. Fig. 10 presents the second field example. The GPR imaging in second scenery [Fig. 10(a)] is more complicated than the first one. As the depth increases, the amplitude of hyperbolic reflections becomes weaker, and some reflected hyperbolic shapes are incomplete because of overlap and intersection with each other. Even under such complex condition, the developed model can still accurately detect and segment the hyperbolae at different depths [Fig. 10(b)], and then visualize multiple mask targets, irrespective of the complex background [Fig. 10(c)]. The fitting results can be found in Fig. 10(d).

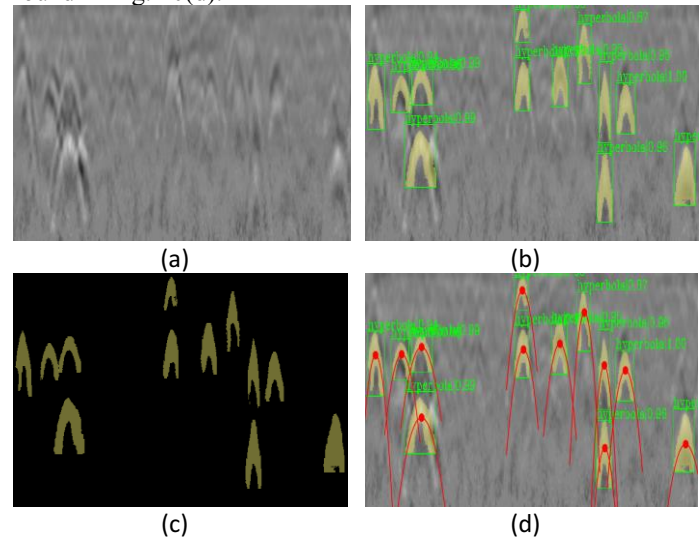


Fig. 10. Second example of GPR root localization results processed by the proposed method. (a) Pre-processed root B-scan. (b) Detected and segmented results. (c) Only picked mask results. (d) Hyperbolic curve fitting with peaks (red point).

More field results are shown in Fig. 11 and Fig. 12. The results demonstrated that the developed model is able to detect and segment object signatures in GPR scans in an automatic and accurate manner.

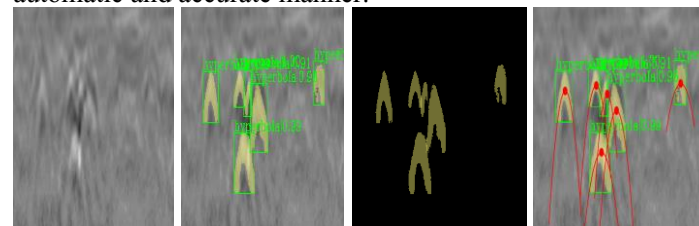


Fig. 11. Third example of GPR root localization results.

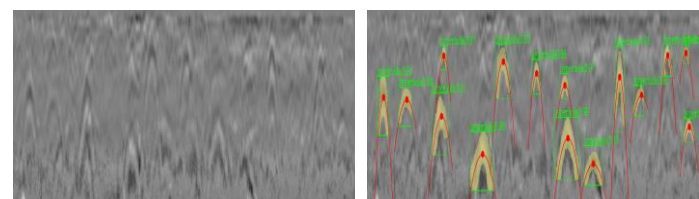


Fig. 12. Fourth example of GPR root localization results.

B. Evaluation Metrics

Average Precision (AP) is used as a criterion for evaluating the performance of the bbox and mask. AP denotes AP at an average of IOU from 0.5 to 0.95 with 0.05 interval. AP_{50} and AP_{75} mark AP at IoU=0.5, IoU=0.75, respectively. AP for objects with different scales: small ($area < 32^2$) and medium ($32^2 < area < 96^2$) are denoted as AP_s and AP_M .

Average Recall (AR) is another metric of RPN results. It averages the recalls of IOU thresholds between 0.5 and 0.95. In particular, AR_{100} , AR_{300} and AR_{1000} are defined to explain AR performance when an image is assigned with 100, 300, and 1000 proposals. For small- (AR_s), medium- (AR_M), or large- (AR_L) scale objects, the AR are computed for 100 proposals. The standard COCO metric is used to evaluate the test results, and it represents the average mAP level when IOUs varies from 0.5 and 0.95.

C. Performance Evaluation

Fig. 13 plots the learning accuracy and loss curves to reflect the performance of the improved MS R-CNN during training period. Accuracy varies with iterations in Fig. 13(a). The x-coordinate and y-coordinate represent iteration and accuracy, respectively. At the beginning of the iteration, its accuracy has climbed to 90%. In subsequent iterations, the accuracy remains above 90% and gradually increases with iterations. It remains stable in 17500-20000 iterations with an accuracy of approximately 97%. Fig. 13(b) shows the distributions of three losses: classification loss, bbox loss, and mask loss. It can be observed all losses remain below about 0.35% after 2500 iterations. The distribution of the classification loss is almost close to the bbox loss, and both are lower than the mask loss.

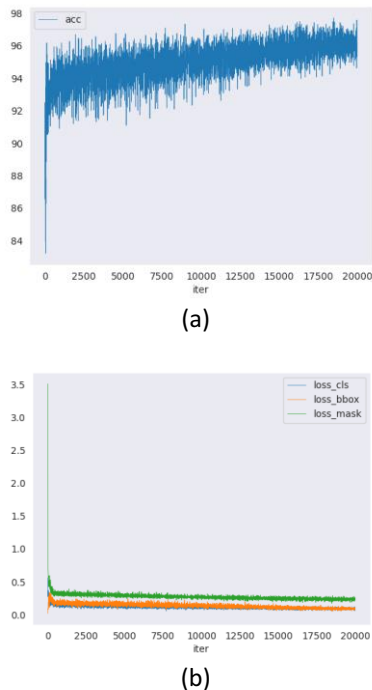


Fig. 13. (a) Accuracy and (b) loss distribution (classification loss, bbox loss, mask loss) of the improved MS R-CNN.

The performance of the improved MS R-CNN is compared with the previous three instance segmentation frameworks from the following three aspects: detection, segmentation, and proposal. The previous frameworks are implemented using the same configuration strategy as our proposed framework. As shown in Table 3, the improved MS R-CNN with customized anchoring scheme and transfer learning achieves 40.0% AP and 65.3% AP_{50} , respectively. Compared to the common MS R-CNN, it increases AP by 3.2%. In addition, in terms of detection results, the common MS R-CNN outperforms Mask R-CNN and Cascade Mask R-CNN [44]. Notably, the enhanced MS R-CNN gains the highest accuracy of 48.7% at AP_s . This is because most root targets present hyperbola signatures with a small size. From Table 4, we can obtain acceptable segmentation results with 35.0% AP and 38.6% AP_{50} . For small and medium objects, the enhanced MS R-CNN outperforms the Mask R-CNN, which demonstrates the efficiency of the enhanced MS R-CNN for objects with scale of $1^2 < area < 96^2$. Table 5 provides AR to evaluate proposal results of the proposed anchoring scheme. The enhanced MS R-CNN obtains 35.1% AR_{100} , 40.2% AR_{300} , and 46.1% AR_{1000} , respectively, which is superior to the other three frameworks. It shows the performance of RPN can be fully utilized by effectively balancing training methods. Since the improved MS R-CNN customizes the anchor boxes with aspect ratios of [0.3492, 0.5406, 0.8528] in RPN stage, it can generate the corresponding proposals to better fit the root target signatures.

TABLE III
COMPARISON OF DETECTION RESULTS VIA DIFFERENT INSTANCE SEGMENTATION FRAMEWORKS (IN PERCENT).

Framework	AP	AP_{50}	AP_{75}	AP_s	AP_M
Mask R-CNN	30.8	53.3	27.5	40.3	18.2
Cascade Mask R-CNN	33.9	65.0	25.8	33.3	22.1
MS R-CNN	36.8	64.6	37.8	37.7	29.7
Improved MS R-CNN	40.0	65.3	37.0	48.7	24.8

TABLE IV
COMPARISON OF SEGMENTATION RESULTS VIA DIFFERENT INSTANCE SEGMENTATION FRAMEWORKS (IN PERCENT).

Framework	AP	AP_{50}	AP_{75}	AP_s	AP_M
Mask R-CNN	27.9	31.4	21.2	27.7	27.8
Cascade Mask R-CNN	34.3	38.2	21.7	32.6	29.5
MS R-CNN	33.6	35.4	26.9	30.3	36.5
Improved MS R-CNN	35.0	38.6	24.5	32.7	41.5

TABLE V
COMPARISON OF PROPOSAL RESULTS VIA DIFFERENT INSTANCE SEGMENTATION FRAMEWORKS (IN PERCENT).

Framework	AR_{100}	AR_{300}	AR_{1000}	AR_s	AR_M
Mask R-CNN	27.0	27.0	28.7	40.0	24.2
Cascade Mask R-CNN	32.6	33.3	37.4	40.1	32.1
MS R-CNN	33.0	35.0	36.7	42.5	31.3
Improved MS R-CNN	35.1	40.2	46.1	52.5	33.1

TABLE VI
RUNNING TIME OF DIFFERENT DEEP FRAMEWORKS.

Framework	Running Time (h:min:s)
Mask R-CNN	5:10:00
Cascade Mask R-CNN	4:41:00
MS R-CNN	5:33:00
Improved MS R-CNN	5:31:00

Table 6 details the running time of all frameworks. By observing the MS R-CNN and the improved MS R-CNN, their training time lasted about five and a half hours, which further demonstrated that the improved MS R-CNN achieves acceptable performance without increasing additional running time.

We evaluated the effectiveness of two components: the customized anchoring scheme and the transfer learning technique, shown in Fig. 14. The common MS R-CNN configured with a single customized anchor component is more effective than that without such configuration. It shows our work is valuable for generating proposals of feature maps in RPN stage by utilizing designed anchors. The single transfer learning component also presents an advantage in the aspect of recall. Both branches are integrated into the common MS R-CNN, which brings obvious improvement to AR_{1000} and AR_s . The enhanced MS R-CNN has the ability to capture information for small scale objects.

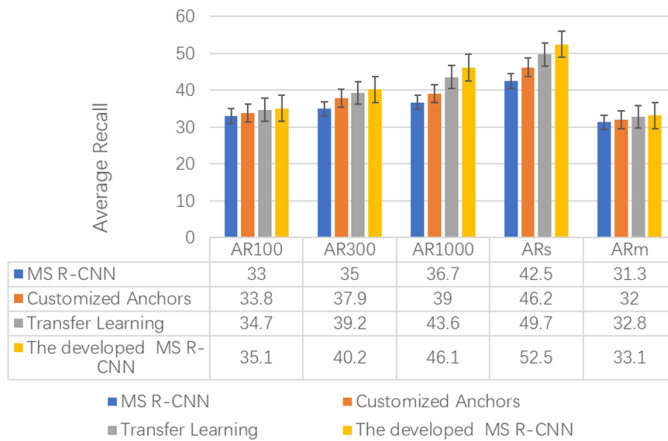


Fig. 14. Effect of each component on the improved MS R-CNN.

D. Comparison of detection result

To further verify the performance of the improved MS R-CNN, an additional comparative experiment is conducted based on the machine learning method. In this work, we perform a Histogram of Oriented Gradients (HOG) feature extraction and train a linear SVM classifier. The code is implemented using Python scikit-learn library based on the open-source vehicle detection toolbox (<https://github.com/hortovanyi/udacity-vehicle-detection-project>). Training a SVM classifier requires both positive samples (with hyperbolas) and negative samples (without hyperbolas). A fixed sample size of 48×48 pixels is used to clipped samples

from the GPR images. A total of 16872 training samples is obtained, including 8436 positive samples and 8436 negative samples (<https://github.com/PouriaAI/GPR-Detection>), as illustrated in Fig. 15. The HOG implementation used in our experiments uses L2-norm normalization. The parameters are set as: orient = 9, and pix_per_cell = 16. Then, we implement a sliding-window technique over GPR image and use the trained classifier to search for hyperbola targets. The sliding window is of size 50×40 pixels and the overlap value is set to 0.7. Based on this technique, many false positive samples and multiple overlapping bboxes are obtained. The heatmap method is adapt to filter redundant bboxes and estimate a single bbox for each hyperbola detected. The count number threshold is set to 4 for heatmap.

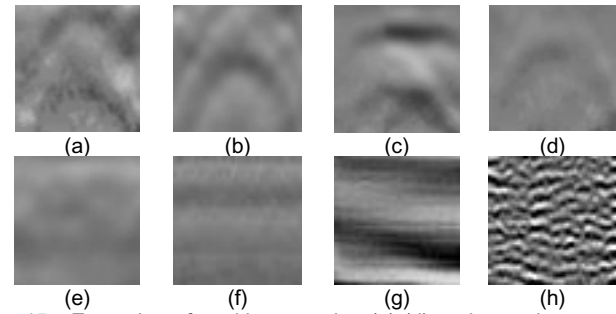


Fig. 15. Examples of positive samples (a)-(d) and negative samples (e)-(h).

To visualize the comparison performance of the above two models, Fig. 16(a) and Fig. 16(b) show the detection results using SVM classifier based on the original GPR images Fig. 8(a) and Fig. 10(a), respectively. The output results in Fig. 8(b) and Fig. 10(b) are generated by the improved MS R-CNN, in which the root hyperbolic signatures can be effectively found. Compared with Fig. 8(b), Fig. 16(a) detects many redundant boxes indicated by green arrow. Compared with Fig. 10(b), Fig. 16(b) omits one hyperbolic target (marked by yellow rectangle), and there are many redundant boxes (marked by green arrows). The comparative results illustrate the performance of the HOG feature-based linear SVM classifier is inferior to the improved MS R-CNN, and it brings many false positive samples and many redundant computations.

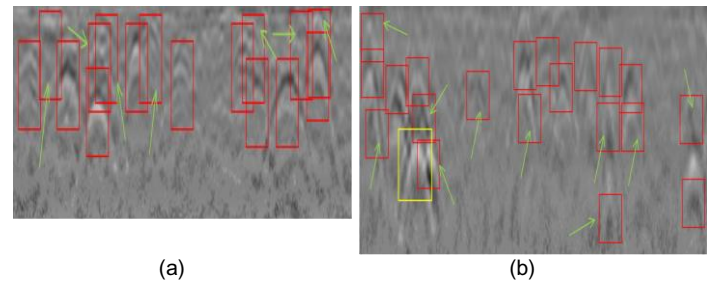


Fig. 16. Detection results obtained by the HOG feature-based linear SVM classifier on (a) original GPR image Fig. 8(a), and (b) original GPR image Fig. 10(a). (Green arrow indicates redundant box and yellow rectangle indicates missing target)

V. CONCLUSION

In this paper, an improved DL model is developed and tested with real GPR data for automatic object signature detection and segmentation. The MS R-CNN is adopted as the main framework. To improve its performance, a novel anchoring scheme is proposed and integrated into MS R-CNN to replace the previous fixed ones (i.e., scale and aspect ratio). This scheme considers the imaging features of underground objects and solves the problem of low detection accuracy of small targets in GPR scans. Considering the need for large datasets for training the DL model, a transfer learning technique is developed to address the issue of lacking training data. Due to the similarity between rebar and tree root signatures in GPR scans, the rebar dataset is picked as the 'source set' and a new model with pre-trained parameters and weights is obtained. The 'target set', root dataset, is exploited to fine-tune the model. The new anchoring scheme and the transfer learning technique improve the performance of the MS R-CNN in bbox detection task, mask segmentation task, and proposal generation task. The experimental results illustrated the feasibility and efficacy of the improved MS R-CNN for automating the signature recognition procedure.

There are some limitations that could be addressed in future work. First, the proposed method requires a large scale of field dataset for training the model. Due to the complexity and unpredictability of field site, it is difficult to obtain GPR datasets. Efforts need to be made to collect large amount of data to improve the performance of the model. Second, DL model needs to be designed to match each type of input characteristics. Due to the non-intuitive and difficult explanation of features extracted by DL model, designing a corresponding model for certain problem is a challenge. Therefore, future research can concentrate on the exploration and interpretation of the internal structure of DL model. Third, the efficacy of this work relied on quality of root dataset. Since the dielectric constant is highly related to soil condition, in the case of high soil water content, the reflected GPR signal is so weak that root targets may be missed. Future research could focus on the signal processing methods to eliminate clutters and enhance weak hyperbola in GPR images. Finally, our recognition about root targets is inferior to the linear object recognition such as rebar or pipe. That is because root system, compared to those linear objects, usually faces more challenges about the identification of root characteristics: (1) root distribution is irregular; (2) root size and root depth are unknown; (3) underground soil condition is complex. These factors do constrain the automatic and real-time development of large-scale root system. In future studies, a deeper understanding towards root itself property is required. In addition, the attention can be paid on the distinctions of the hyperbolic signature between root and non-root targets.

REFERENCES

1. Daniels, D.J., *Ground penetrating radar*. Encyclopedia of RF and Microwave Engineering, 2005.
2. Tong, et al., *Innovative method for recognizing subgrade defects based on a convolutional neural network*. CONSTRUCTION AND BUILDING MATERIALS, 2018.
3. Tong, Z., et al., *Pavement-distress detection using ground-penetrating radar and network in networks*. Construction and Building Materials, 2019. **233**: p. 117352.
4. Gao, J., et al., *Autonomous pavement distress detection using ground penetrating radar and region-based deep learning*. Measurement, 2020. **164**: p. 108077.
5. Bianchini Ciampoli, L., et al., *Signal Processing of GPR Data for Road Surveys*. Geosciences, 2019. **9**(2).
6. Gustavsen, L., et al., *The Effectiveness of Large-Scale, High-Resolution Ground-Penetrating Radar Surveys and Trial Trenching for Archaeological Site Evaluations—A Comparative Study from Two Sites in Norway*. Remote Sensing, 2020. **12**(9): p. 1408.
7. Cozzolino, M., et al., *Imaging Buried Archaeological Features through Ground Penetrating Radar: The Case of the Ancient Saepinum (Campobasso, Italy)*. Geosciences, 2020. **10**(6): p. 225.
8. Zhang, X., et al., *Evaluation of a ground penetrating radar to map the root architecture of HLB-infected citrus trees*. Agronomy, 2019. **9**(7): p. 354.
9. Isaac, M.E. and L.C. Anglaaere, *An in situ approach to detect tree root ecology: linking ground - penetrating radar imaging to isotope - derived water acquisition zones*. Ecology and evolution, 2013. **3**(5): p. 1330-1339.
10. Cui, X., et al., *Estimating tree-root biomass in different depths using ground-penetrating radar: Evidence from a controlled experiment*. IEEE Transactions on Geoscience and Remote Sensing, 2012. **51**(6): p. 3410-3423.
11. Guo, L., et al., *Application of ground penetrating radar for coarse root detection and quantification: a review*. Plant and soil, 2013. **362**(1-2): p. 1-23.
12. Wu, Y., et al., *Ground-penetrating radar-based automatic reconstruction of three-dimensional coarse root system architecture*. Plant and soil, 2014. **383**(1-2): p. 155-172.
13. Prasanna, P., et al., *Automated Crack Detection on Concrete Bridges*. IEEE Transactions on Automation and Engineering, 2016. **13**(2): p. 591-599.
14. Okazaki, Y., et al., *Applicability of machine learning to a crack model in concrete bridges*. Computer - Aided Civil and Infrastructure Engineering, 2020.
15. Li, S., H. Cai, and V.R. Kamat, *Uncertainty-aware geospatial system for mapping and visualizing underground utilities*. Automation in Construction, 2015. **53**(may): p. 105-119.
16. Sagnard, F., et al., *Utility detection and positioning on the urban site Sense-City using Ground-Penetrating Radar systems*. Measurement, 2016. **88**: p. 318-330.
17. Li, W., et al., *Tree root automatic recognition in ground penetrating radar profiles based on randomized Hough transform*. Remote Sensing, 2016. **8**(5): p. 430.
18. Wang, M., J. Wen, and W. Li, *Qualitative Research: The Impact of Root Orientation on Coarse Roots Detection Using Ground-Penetrating Radar (GPR)*. BioResources, 2020. **15**(2): p. 2237-2257.
19. Maas, C. and J.R. Schmalzl, *Using pattern recognition to automatically localize reflection hyperbolas in data from ground penetrating radar*. Computers & Geosciences, 2013. **58**: p. 116-125.
20. D, Y.W.A.B.C., G.C.A.B. C, and J.X. E, *Semi-automatic detection of buried rebar in GPR data using a genetic algorithm*. Automation in Construction, 2014. **114**.
21. Harkat, H., et al., *GPR target detection using a neural network classifier designed by a multi-objective genetic algorithm*. Applied Soft Computing, 2019.
22. Kaur, P., et al., *Automated GPR Rebar Analysis for Robotic Bridge Deck Evaluation*. IEEE Transactions on Cybernetics, 2015. **46**(10): p. 2265-2276.
23. Pham, M.-T. and S. Lefèvre, *Buried object detection from B-scan ground penetrating radar data using Faster-RCNN*. in IGARSS 2018-2018 IEEE International Geoscience and Remote Sensing Symposium. 2018. IEEE.

24. Lei, W., et al., *Automatic hyperbola detection and fitting in GPR B-scan image*. Automation in Construction, 2019. **106**: p. 102839.
25. Man-Sung Kang, Y.-K.A., *Frequency-Wavenumber Analysis of Deep Learning-based Super Resolution 3D GPR Images*. Remote Sensing, 2020. **12**(18): p. 3056.
26. Dinh, K., N. Gucunski, and T.H. Duong, *An algorithm for automatic localization and detection of rebars from GPR data of concrete bridge decks*. Automation in Construction, 2018. **89**(MAY): p. 292–298.
27. Zhang, J., et al., *Automatic detection of moisture damages in asphalt pavements from GPR data with deep CNN and IRS method*. Automation in Construction, 2020. **113**.
28. He, K., et al. *Deep residual learning for image recognition*. in *Proceedings of the IEEE conference on computer vision and pattern recognition*. 2016.
29. Redmon, J. and A. Farhadi. *YOLO9000: better, faster, stronger*. in *Proceedings of the IEEE conference on computer vision and pattern recognition*. 2017.
30. Ren, S., et al. *Faster r-cnn: Towards real-time object detection with region proposal networks*. in *Advances in neural information processing systems*. 2015.
31. Hou, F., et al., *Improved Mask R-CNN with Distance Guided Intersection over Union for GPR Signature Detection and Segmentation*. Automation in Construction, 2020. **121**.
32. He, K., et al. *Mask r-cnn*. in *Proceedings of the IEEE international conference on computer vision*. 2017.
33. Zhou, X., H. Chen, and J. Li, *An automatic GPR B-scan image interpreting model*. IEEE Transactions on Geoscience and Remote Sensing, 2018. **56**(6): p. 3398-3412.
34. Dou, Q., et al., *Real-time hyperbola recognition and fitting in GPR data*. IEEE Transactions on Geoscience and Remote Sensing, 2016. **55**(1): p. 51-62.
35. Huang, Z., et al. *Mask scoring r-cnn*. in *Proceedings of the IEEE conference on computer vision and pattern recognition*. 2019.
36. Liu, W., et al. *Ssd: Single shot multibox detector*. in *European conference on computer vision*. 2016. Springer.
37. Lin, T.-Y., et al. *Focal loss for dense object detection*. in *Proceedings of the IEEE international conference on computer vision*. 2017.
38. Redmon, J. and A. Farhadi, *Yolov3: An incremental improvement*. arXiv preprint arXiv:1804.02767, 2018.
39. Krishna, K. and M.N. Murty, *Genetic K-means algorithm*. IEEE Transactions on Systems, Man, and Cybernetics, Part B (Cybernetics), 1999. **29**(3): p. 433-439.
40. Weisstein, E.W., *Hyperbola*. <https://mathworld.wolfram.com/>, 2003.
41. Bralich, J., et al. *Improving convolutional neural networks for buried target detection in ground penetrating radar using transfer learning via pretraining*. in *Detection and Sensing of Mines, Explosive Objects, and Obscured Targets XXII*. 2017. International Society for Optics and Photonics.
42. Xie, S., et al. *Aggregated residual transformations for deep neural networks*. in *Proceedings of the IEEE conference on computer vision and pattern recognition*. 2017.
43. Liu, Z., et al. *Using adaptive cross approximation to accelerate simulation of B-scan GPR for detecting underground pipes*. in *2016 Progress in Electromagnetic Research Symposium (PIERS)*. 2016. IEEE.
44. Cai, Z. and N. Vasconcelos, *Cascade R-CNN: high quality object detection and instance segmentation*. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2019.



Feifei Hou received the B.S. degree in communication engineering from Henan Polytechnic University (HPU), China, in 2015. She studied as a visiting scholar in the Department of Civil and Environmental Engineering at the University of Tennessee, Knoxville, USA from 2019 to 2020. She now is a candidate Ph.D. in computer science, Central South University (CSU), Changsha, China. Her research interests include computer vision, deep learning, machine learning, ground penetrating radar, and civil infrastructure applications.



Wentai Lei received the B.S. degree in communication & navigation engineering at 2000 and M.S. degree in radar engineering at 2003, both from Air Force Engineering University, China. He received the Ph.D. degree in National University of Defense Technology, China in 2006. He is an associate professor in Central South University now. From 2009 to 2011, he conducted postdoctoral research in school of geoscience and information physics, Central South University.

His research interest includes gesture radar signal feature extraction and recognition technique by using machine learning, ground penetrating cognitive radar technique with soft-defined radio implementation and electromagnetic inversion scattering imaging technique by using physical optical approximation.



Shuai Li is an assistant professor in the Department of Civil and Environmental Engineering at the University of Tennessee, Knoxville. He graduated from Purdue University with a Ph.D. degree in Civil Engineering, a master's degree in industrial engineering, and a master's degree in economics. He conducts fundamental research in sensing, automation, robotics, and visualization, and applies the techniques in numerous applications, including smart construction, disaster response, and manufacturing. His e-mail address is sl48@utk.edu.



Jingchun Xi received the B.S. degree in communication engineering, Beijing University of Chemical Technology, Beijing, China, in 2017. He received the M.S. degree in electronic communication engineering, Central South University (CSU), Changsha, China, in 2020. His research interests include computer vision, signal processing, and ground penetrating radar.