

# Learning to Remove Clutter in Real-World GPR Images Using Hybrid Data

Hai-Han Sun, Weixia Cheng, and Zheng Fan

**Abstract**—The clutter in the ground-penetrating radar (GPR) radargram disguises or distorts subsurface target responses, which severely affects the accuracy of target detection and identification. Existing clutter removal methods either leave residual clutter or deform target responses when facing complex and irregular clutter in the real-world radargram. To tackle the challenge of clutter removal in real scenarios, a clutter-removal neural network (CR-Net) trained on a large-scale hybrid dataset is presented in this study. The CR-Net integrates residual dense blocks into the U-Net architecture to enhance its capability in clutter suppression and target reflection restoration. The combination of the mean absolute error (MAE) loss and the multi-scale structural similarity (MS-SSIM) loss is used to effectively drive the optimization of the network. To train the proposed CR-Net to remove complex and diverse clutter in real-world radargrams, the first large-scale hybrid dataset named CLT-GPR dataset containing clutter collected by different GPR systems in multiple scenarios is built. The CLT-GPR dataset significantly improves the generalizability of the network to remove clutter in real-world GPR radargrams. Extensive experimental results demonstrate that the CR-Net achieves superior performance over existing methods in removing clutter and restoring target responses in diverse real-world scenarios. Moreover, the CR-Net with its end-to-end design does not require manual parameter tuning, making it highly suitable for automatically producing clutter-free radargrams in GPR applications. The CLT-GPR dataset and the code implemented in the paper can be found at <https://haihan-sun.github.io/GPR.html>.

**Index Terms**—Clutter removal, deep learning, ground-penetrating radar, convolutional neural network.

## I. INTRODUCTION

GROUND-PENETRATING radar (GPR) has been a powerful tool for detecting subsurface targets in many fields, such as civil engineering, geology engineering, and military [1], [2]. However, the strong background clutter in the GPR radargram caused by the reflection from the air-to-ground interface and the heterogeneity of subsurface materials disguises or obscures the target response, especially when the target is at a shallow depth. Wideband GPR systems have been implemented [3]–[8] and high-resolution algorithms have been developed [9]–[15] to enhance discrimination between targets and clutter. Since the presence of environmental clutter in the radargram significantly challenges the accuracy of target detection and interpretation, it is imperative to eliminate environmental clutter and restore target responses, which is the focus of this work.

Over the years, many methods have been investigated to suppress clutter in GPR radargrams. Conventional methods include mean subtraction methods [16]–[18] and subspace-based methods such as singular value decomposition (SVD)

[19]–[21], principal component analysis (PCA), and independent component analysis (ICA) [22], [23]. The mean subtraction methods regard the averaged trace of raw traces in a selected window in a B-scan as the background clutter trace, and subtract it from raw traces to eliminate clutter. They perform well in homogeneous subsurface environments with flat surfaces where the averaged trace in no-object regions fully captures the background clutter information. However, their performance is degraded in heterogeneous environments with rough surfaces where the background clutter varies from trace to trace in the B-scan [21]. The subspace-based methods project the original B-scan matrix onto the clutter and target components based on the difference in signal strength, and remove the clutter components from the raw data to mitigate clutter. They assume that the clutter signal is much stronger than the target response and thus the clutter information is contained in the dominant components. However, the clutter and target components cannot be well separated when their signals have similar strengths. As a result, the target response could be deteriorated by the clutter removal operation using these methods [24].

Based on the prior that the clutter caused by antennas' crosstalk and surface reflection is relatively constant so it has a low-rank characteristic, whereas the target response has a sparse property, non-negative matrix factorization (NMF) [25] is proposed to extract the low-rank components and remove them from the raw data to suppress clutter. Following that, the methods based on low-rank and sparse decomposition (LRSD) are investigated, including the robust PCA (RPCA) [26], [27], go decomposition (GoDec) [28], [29], and robust NMF (RNMF) [30]. These methods separate the target response and the clutter by optimizing a low-rank and sparse matrix representation problem. However, the target response that has wide and flat hyperbolic features can degrade the sparsity and result in low effectiveness of these methods [27]. Moreover, the performance of these methods highly depends on the selection of hyperparameters that control the trade-off between the low-rank components and sparse components, and the suitable hyperparameters vary from case to case [30]. Therefore, these methods may have inconsistent clutter removal performance under different subsurface scenarios. Similarly, the morphological component analysis (MCA) [31] and dictionary learning-based method [32] are presented to separate the clutter and target responses using their respective sparse representations or dictionaries. The target responses are subsequently reconstructed using the target dictionaries. The clutter removal performance of these methods highly depends on the selection of accurate dictionaries [33]. Considering

the limitations of the conventional methods, a more effective solution is desired to improve the clutter removal performance in diverse real-world subsurface environments.

With the strong feature representation and learning capability, deep learning-based methods have been involved in solving challenging GPR tasks, such as target detection [34], [35], characterization [36]–[38], and inverse imaging [39], [40]. There is a rising trend to apply neural networks in the clutter removal task. In [41], a robust autoencoder-based method is introduced to solve the low-rank and sparse matrix representation problem. This method achieves better performance than RPCA [26] as the autoencoder offers nonlinear solutions. In [42], a network model constructed with 1D convolutional layers is presented to separate target responses from clutter in GPR A-scans. By using low-rank components and sparse components decomposed by RNMF as guidance, the network achieves similar performance with RNMF without tuning hyperparameters. In [33], a network model based on the convolutional autoencoder (CAE) is presented to learn the clutter removal in 2D B-scans using a synthetic dataset. The network outperforms most of the existing methods in multiple scenarios, but its clutter removal capability is degraded in real-world GPR radargrams. This is because the network is only trained on the synthetic data that cannot capture the complex and diverse real-world clutter distributions.

Previous deep learning-based methods have shown promising capability in removing GPR clutter, but their generalizability for real-world scenarios is highly constrained by the synthetic training data and the simple network structure. This work devotes to address the challenges of deep learning-based methods in removing clutter in real-world scenarios while enhancing the clutter removal capability in general by making three contributions:

- In view of the limitations of using synthetic data to train the data-driven clutter removal method, we built the first large-scale hybrid dataset with real-world clutter distributions named CLT-GPR dataset. The proposed dataset contains diverse clutter distributions from both the synthetic data and measured data collected using different GPR systems in different environments. Experimental results verify that the dataset serving as training data could significantly boost the generalization capability of the data-driven method in eliminating clutter in real-world radargrams.
- Different from previous deep learning-based methods that use simple network architectures, we leverage the powerful learning capability of deep neural network and present a clutter-removal network (CR-Net). The CR-Net

is featured by the integration of residual dense blocks (RDBs) into the U-Net architecture. The network extracts the hierarchical features of the input raw radargram with different receptive fields, retains the features related to the target reflection and suppresses the unwanted clutter features, and finally reconstructs the clutter-free radargram. The inclusion of RDBs enhances the network capability in adaptively removing clutter without deteriorating target responses.

- The loss function that drives network optimization has a critical impact on the network performance. In this work, instead of using conventionally adopted mean squared error (MSE) loss, we take one step forward by investigating the effects of different losses on the network performance for the clutter removal task. It is experimentally proved that the combined loss function of the mean absolute error (MAE) loss and the multi-scale structure similarity (MS-SSIM) loss leads to greatly improved clutter removal performance than the MSE loss.

Extensive experiments demonstrate that the built CLT-GPR dataset, the design of the CR-Net, and the combined loss function benefit from each other, producing a network with superior and consistent performance in suppressing clutter in diverse scenarios. The well-trained CR-Net automatically removes clutter and restores the target response without any selection of hyperparameters. Moreover, several case studies verify that our method outperforms the existing methods by a large margin in terms of both quantitative metrics and visual quality.

The rest of the paper is arranged as follows. Section II describes the collection of the CLT-GPR dataset. Section III presents the CR-Net architecture and the combined loss function that drive the network optimization. In Section IV, extensive experiments are conducted to evaluate the performance of the CR-Net and demonstrate the advantages of our method in various scenarios. Finally, conclusion is drawn in Section V.

## II. CLT-GPR DATASET COLLECTION

As the existing dataset for the clutter removal task only contains synthetic data [33], [42], it cannot allow data-driven methods to learn complex and diverse clutter distributions in real-world radargrams, thus constraining the generalization capability of the data-driven methods in real scenarios. To address this issue, we built the CLT-GPR dataset containing diverse real-world clutter for training data-driven clutter removal methods. The CLT-GPR dataset is composed of three

TABLE I  
PROPERTIES OF THE THREE SUB-DATASETS IN THE CLT-GPR DATASET

| Sub-dataset           | GPR system                                                                           | Subsurface environment                                                 | Number of data |
|-----------------------|--------------------------------------------------------------------------------------|------------------------------------------------------------------------|----------------|
| Synthetic sub-dataset | 1.5-GHz GSSI antenna in gprMax [43]–[46]                                             | Six different soil conditions with four types of soil surfaces         | 1,920          |
| Sand sub-dataset      | Multi-polarimetric GPR system                                                        | Sandy soil with uneven surface and random distributed moisture content | 6,000          |
| Concrete sub-dataset  | Single-polarimetric GPR system with mono-static and bi-static antenna configurations | Concrete                                                               | 4,000          |

sub-datasets: a synthetic sub-dataset, a sand sub-dataset, and a concrete sub-dataset. The characteristics of each sub-dataset are listed in Table I. The synthetic sub-dataset includes simulated clutter in different soil and surface conditions, which is detailed in subsection A. The sand sub-dataset includes clutter collected by a multi-polarimetric GPR system in a sandy field, which is described in subsection B. The concrete sub-dataset contains clutter collected by a single-polarized GPR system with mono-static and bi-static antenna configurations in concrete environment, which is presented in subsection C. The different GPR systems and subsurface scenarios in the dataset preparation provide diverse distributions of the real-world clutter. Each pair of data in the dataset includes a raw B-scan radargram and its corresponding clutter-free image, which facilitates the supervised learning of data-driven methods. The preparation details of these sub-datasets are described as follows.

#### A. Synthetic Sub-Dataset

A synthetic GPR dataset is generated using the GPU-accelerated open-source software gprMax [43], [44]. The simulation scenario is shown in Fig. 1(a). The simulation domain covers an area of  $100 \times 15 \times 40 \text{ cm}^3$ . The absorbing boundary condition is applied to reduce the side reflection. The spatial discretization is 0.2 cm in the  $x$ -,  $y$ - and  $z$ -directions. The build-in GSSI antenna operating at 1.5 GHz is used as the transmitter and receiver [45], [46]. It has a size of  $17 \times 10.8 \times 4.5 \text{ cm}^3$  and is located 5 cm above the soil surface. The antenna is moved along the  $x$ -direction with a step of 1 cm. 80 A-scans are collected along the scan traces to generate a B-scan.

To generate a synthetic dataset covering diverse subsurface scenarios, four types of surfaces, six types of soil, two types of objects, and three numbers of subsurface objects with different depths and radii are modeled in the simulation.

The types of surfaces cover common scenarios, including the flat surface, the grass surface, the rough surface, and the rough surface with surface water. For the grass surface, 10,000 blades of grass are randomly distributed on a flat surface. The height of the blades of grass and the depth of the grass roots both vary from 0 to 4 cm. For the rough surface, the roughness is randomly distributed with a fluctuation range of 4 cm. For the rough surface with surface water, the surface roughness is the same as the rough surface cases, but surface water is added with a depth of 1 cm.

The six types of soil include dry sand, damp sand, dry clay soil, wet clay soil, dry loam soil, and heterogeneous soil. The material properties of different soil conditions (except the heterogeneous soil) are listed in Table II. The heterogeneous soil is composed of 50 different soil fractures with a sand fraction of 0.5, a clay fraction of 0.5, a soil bulk density of  $2 \text{ g/cm}^3$ , a sand density of  $2.66 \text{ g/cm}^3$ , a water volumetric fraction of 0.1%–25%, and fractal dimension of 1.5.

The objects are cylinders made of the perfect electric conductor (PEC) or polyvinyl chloride (PVC). The material properties of the modeled PVC are also listed in Table II. In each combination of soil and surface scenario, cylindrical objects with their radii vary from 1 cm to 5 cm are buried

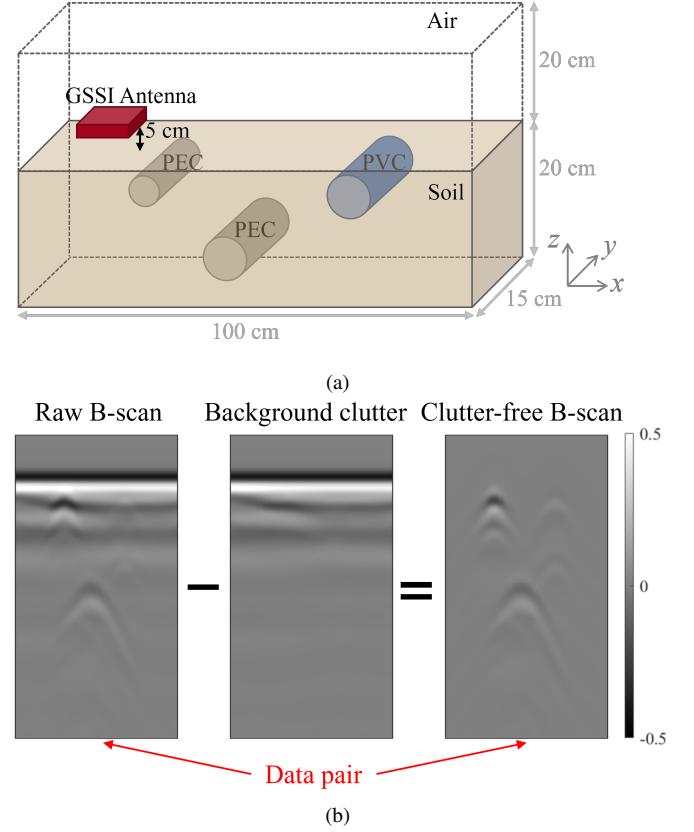


Fig. 1. (a) Illustration of the simulation scenario for the preparation of the synthetic sub-dataset. The 1.5-GHz GSSI antenna is used as the transmitter and receiver to collect GPR signals. Four types of soil surfaces and six types of soil are included as soil conditions to generate diverse subsurface environments. Two types of cylindrical objects (i.e., PVC and PEC) with different radii are included as subsurface targets, and they are placed at different positions. (b) Illustration of the raw B-scan, the corresponding background clutter, and the clutter-free B-scan. The clutter-free B-scan is obtained by subtracting the background clutter from the raw B-scan.

along the  $y$ -direction. The number of subsurface objects varies from one to three. For the one-object case, 20 B-scans are generated, in which 10 B-scans for PVC objects and 10 B-scans for PEC objects. The object is located at the horizontal position of 50 cm but with depths varying from 1 cm to 10 cm. For the two-object and three-object cases, 30 B-scans are generated for each case. The materials of the objects are randomly selected from PEC and PVC, and the horizontal positions and cover depths are randomly selected in the range of 20–80 cm and 0–10 cm, respectively. Special care is taken to avoid the overlapping of multiple objects. A total of 80 B-scans are generated for each combination of soil and surface scenario.

In total, four types of surfaces, six types of soil, and 80 different object arrangements in each soil and surface scenario produces 1,920 raw B-scans. Their corresponding clutter-only background B-scans are obtained by maintaining the same soil condition but without burying objects. The clutter-free B-scans are obtained by subtracting the background B-scans from the raw B-scans. An example of the raw, background, and clutter-free B-scans of a three-object case is shown in Fig. 1(b) to illustrate the process of obtaining a pair of data.

TABLE II  
ELECTROMAGNETIC PROPERTIES OF THE MATERIALS FOR THE SYNTHETIC SUB-DATASET

| Material      | Relative permittivity | Conductivity (S/m) |
|---------------|-----------------------|--------------------|
| Dry sand      | 3.0                   | 0.001              |
| Damp sand     | 8.0                   | 0.01               |
| Dry clay soil | 10.0                  | 0.01               |
| Wet clay soil | 12.0                  | 0.01               |
| Dry loam soil | 10.0                  | 0.001              |
| PVC           | 3.5                   | 0                  |

### B. Sand Sub-Dataset

As the ideal clutter-free image cannot be obtained from the raw radargram in the field experiment, we adopt an alternative method to generate a pair of raw B-scan and its corresponding clutter-free B-scan. We first collect clutter-only radargrams in a controlled environment, and then combine them with the simulated clutter-free images as presented in Section II.A to produce hybrid raw GPR images. These hybrid raw images and their corresponding clutter-free images form the experimental dataset.

In the first field experiment, a multi-polarimetric GPR system is employed to collect real clutter in sandy soil without any subsurface objects, as shown in Fig. 2(a). The sandy field has an uneven surface and randomly distributed moisture levels in the measurement. The height variation of the soil surface is within 4 cm. Different moisture levels are created by adding water to different portions of the soil. The resultant relative permittivity of the soil varies from 3 to 8. These settings are created to increase clutter diversity. The dual-polarized Vivaldi antennas (DPA) operating from 1.3 GHz to 6.0 GHz are used as the transmitter and receiver in a mono-static setup. The antenna is sealed in a foam box and is surrounded by absorbers to reduce noise from the surrounding environment. The antenna has two ports 1 and 2 that excite the horizontal-polarized (H-pol) and vertical-polarized (V-pol) radiation, respectively. The two ports are connected to a vector network analyzer (Keysight VNA P5022A), forming a multi-polarimetric stepped-frequency GPR system. The system simultaneously collects three A-scans of the three polarimetric components, including co-polarized components  $S_{HH}$  and  $S_{VV}$  and cross-polarized component  $S_{HV}$ . The first and second subscripts of the scattering components denote the polarization of receiving and transmitting antennas. For each A-scan, 1,001 samples in a frequency band from 1.3 GHz to 6.0 GHz are recorded. The collected frequency domain data are transformed to the time domain via inverse Fourier transform. For each polarimetric component, 64 A-scans are collected along a scan trace with a step size of 1 cm and are combined into a clutter-only B-scan.

In the experiment, 300 clutter-only B-scans of different polarimetric components are collected. To further increase the clutter diversity, four different frequency bands are selected to process the acquired frequency domain data, which are 1.3 - 3.3 GHz, 1.3 - 4.3 GHz, 1.3 - 5.3 GHz, and 1.3 - 6.0 GHz, respectively. A total of 1,200 clutter-only B-scans are obtained. Each radargram is resized to 256×64 and combined with five randomly selected simulated clutter-free radargrams,

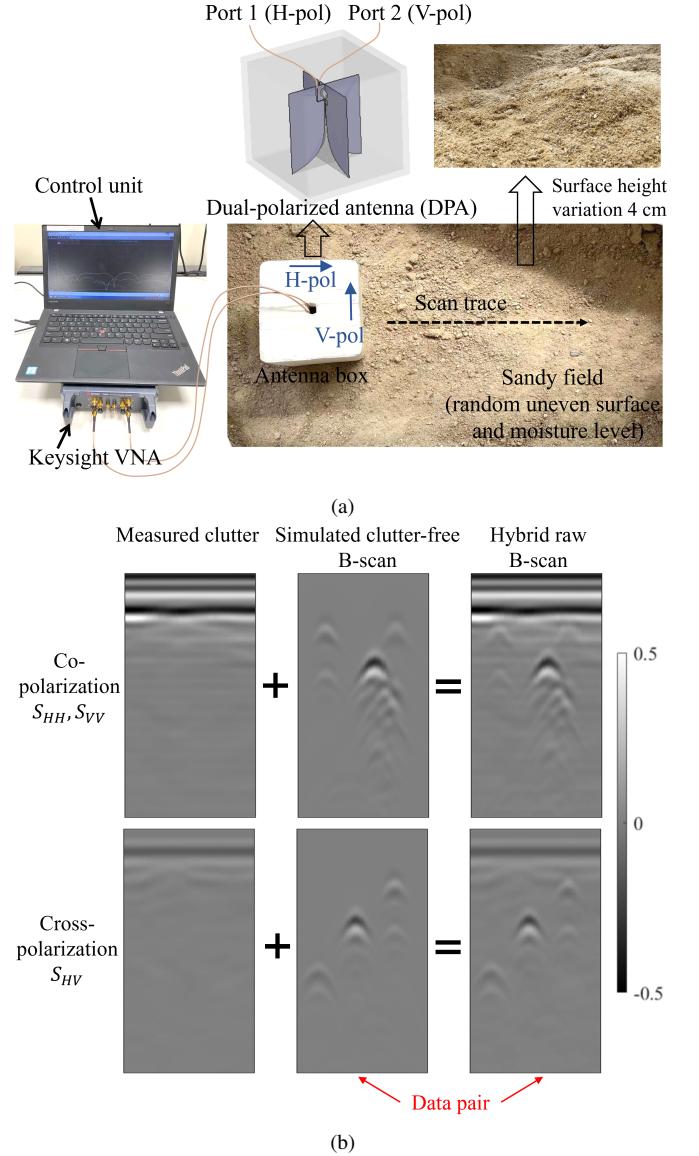


Fig. 2. (a) Demonstration of the measurement scenario using a multi-polarimetric GPR system to collect real clutter data in a controlled sandy field without any subsurface objects. The sandy field has an uneven surface and random distributed moisture levels. (b) Illustration of the collected clutter-only radargram of the co-polarized components ( $S_{HH}$  and  $S_{VV}$ ), and the cross-polarized components ( $S_{HV}$ ), and the process to generate the hybrid B-scan by combining the measured clutter with a simulated clutter-free radargram.

forming 6,000 hybrid data. The clutter-free images are also resized to 256×64 to ease the combination process. The hybrid data and their corresponding clutter-free images form the sand sub-dataset. It is noted that the simulated target response in the clutter-free image was obtained using the 1.5 GHz GPR and therefore may not exactly match the signal response in other frequency bands. However, the objective of the hybrid dataset is to allow the data-driven methods to capture the features of real clutter and remove them, where the presence of the target response only provides supplementary help for the methods to distinguish the characteristics of the clutter from the target response. The operation to produce the hybrid dataset is feasible for the clutter removal task, as demonstrated

by subsequent experiments with the dataset.

Fig. 2(b) illustrates the process of generating the sand sub-dataset. As the soil surface is uneven and the sandy field is heterogeneous, unevenness of the ground reflection and sub-surface environmental clutter can be observed in the collected clutter radargram. In addition, it can be observed that the clutter collected by the co-polarized component and the cross-polarized component are significantly different. The cross-polarized component receives much less antenna crosstalk and surface reflection clutter than the co-polarized components due to polarization mismatch [47]. The clutter collected by different polarimetric components of the GPR system further increases the clutter diversity in the sand sub-dataset. After combining the clutter radargram with the simulated clutter-free radargram, the background clutter interferes and disguises the target responses, which imitates the radargram collected in real scenarios. Therefore, the obtained hybrid radargram contains diverse clutter distributions and resembles the radargrams in real cases.

### C. Concrete Sub-Dataset

In the second field experiment, a single-polarimetric GPR system with both mono-static and bi-static antenna configurations is used to collect clutter data in concrete environments, as shown in Fig. 3(a). The system setup is similar to the one described in Section II.B, except that the equipped antenna is the single-polarized Vivaldi antenna (SPA). Both the mono-static and bi-static antenna configurations are implemented as they provide different clutter coming from the antenna and surface reflection, which enhances the clutter diversity. For each A-scan, the GPR system records 1001 frequency samples from 1.3 GHz to 6.0 GHz. 64 A-scans are collected and are combined into a B-scan.

In the experiment, 200 clutter-only B-scans are collected on 15 concrete samples. Four different frequency bands are selected to process the collected data to further increase the clutter diversity, which are 1.3 - 3.3 GHz, 1.3 - 4.3 GHz, 1.3 - 5.3 GHz, and 1.3 - 6.0 GHz, respectively. The data are subsequently converted into the time domain via Fourier transform. In total, 800 clutter-only B-scans are obtained. Each radargram is resized and combined with five randomly selected clutter-free radargrams, forming 4,000 hybrid data. The hybrid data and their corresponding clutter-free radargrams form the concrete sub-dataset.

Fig. 3(b) shows examples of clutter radargrams collected by the mono-static and bi-static antenna configurations and the obtained hybrid images in the concrete environment. The mono-static setup produces more clutter than the bi-static setup due to antenna self-reflection and ringing effect. As the concrete has a relatively flat surface, the surface clutter presents as horizontal bandings. The surface clutter could severely disguise the target signal in the hybrid data, especially for shallowly located objects, which imitates radargrams of concrete buildings with shallowly buried rebars.

The CLT-GPR dataset consisting of the synthetic sub-dataset and the two hybrid sub-datasets contains diverse clutter distributions of different GPR systems in multiple scenarios.

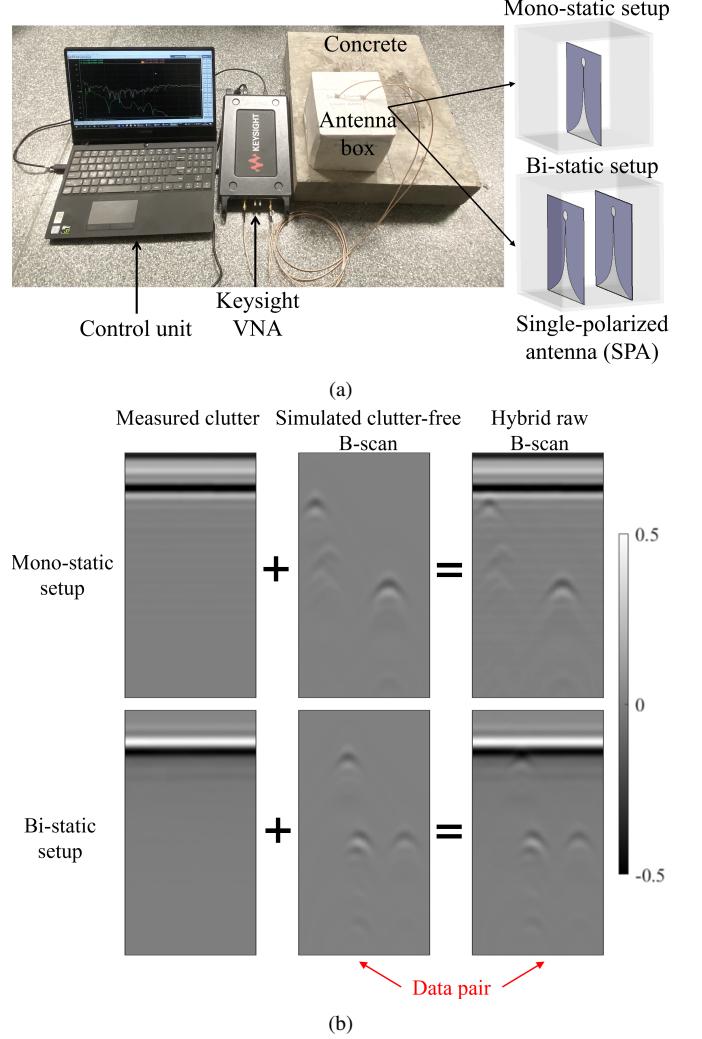


Fig. 3. (a) Demonstration of the measurement scenario using a stepped-frequency GPR system with single-polarized Vivaldi antenna (SPA) in mono-static and bi-static setups to collect real clutter data. The data are collected on rebar-free concrete samples. (b) Illustration of the collected clutter radargrams by the mono-static and bi-static setups and the process to obtain hybrid B-scans.

Therefore, such a large-scale and diverse dataset could provide powerful support for training deep neural networks to remove clutter and restore target responses in real-world scenes. We believe the release of the CLT-GPR dataset could facilitate the development of novel clutter removal methods.

## III. CLUTTER-REMOVAL NEURAL NETWORK

With the collection of CLT-GPR dataset, a clutter-removal neural network, named CR-Net, is designed to eliminate clutter while restoring target responses in radargrams. The network structure and the loss function are described as follows.

### A. Network Architecture

The network architecture of the CR-Net is illustrated in Fig. 4. It consists of an encoder, a decoder, and skip connections between them. The RDBs [48] are integrated into the skip connection paths to adaptively preserve the target response information and reduce the clutter features.

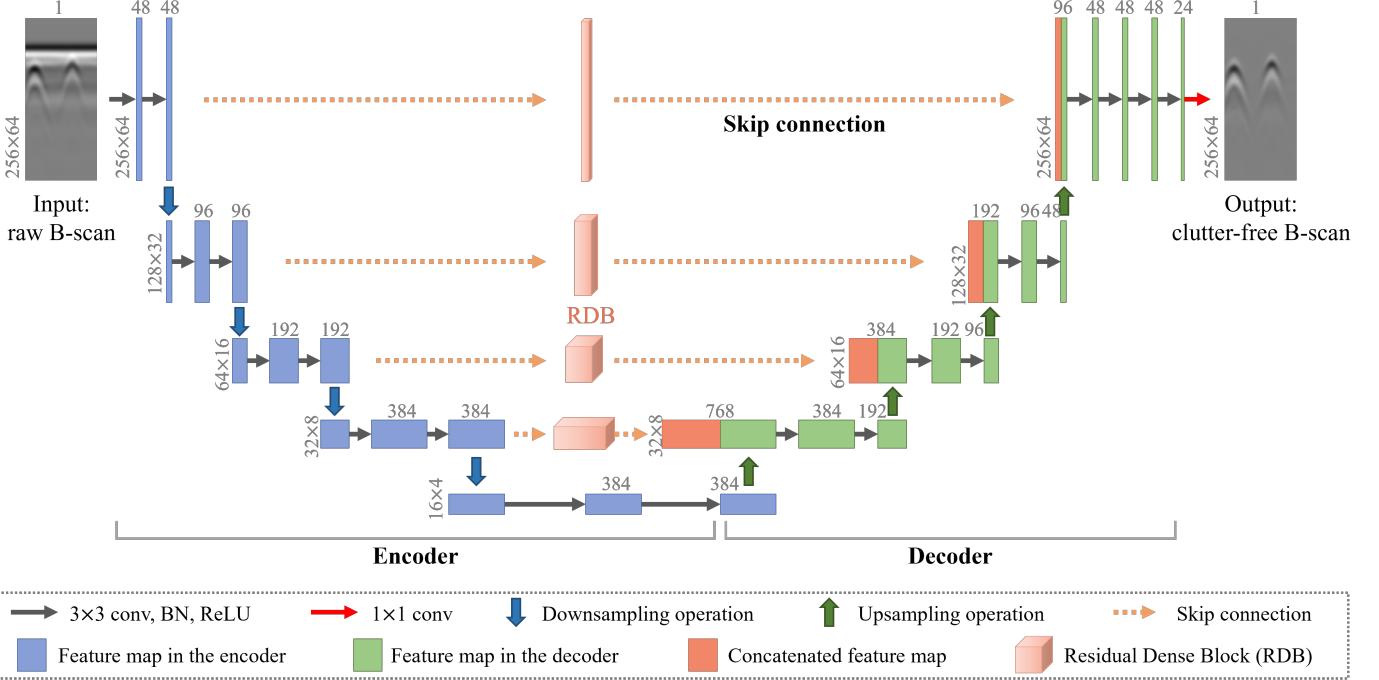


Fig. 4. Overview framework of the proposed CR-Net. It is composed of an encoder, a decoder, and skip connections between them. Residual dense blocks (RDBs) are integrated in the skip connection paths to adaptively preserve features related to object reflection and suppress features of clutter. The number of channels and a size of a feature map are denoted on the top and left of the box that representing a feature map, respectively.

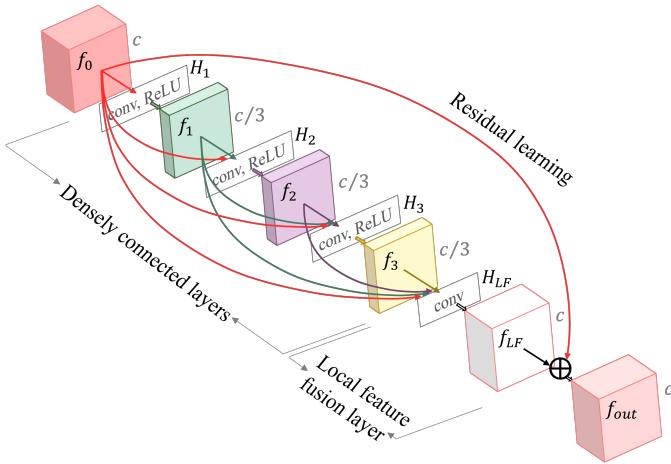


Fig. 5. Illustration of the residual dense block. It includes three densely connected layers, one local feature fusion layer, and one residual learning operation. The main purpose of the block is to adaptively preserve features related to target reflection and reduce the clutter features.  $f$ ,  $c$ , and  $H$  denote the feature map, the channel number of a feature map, and the non-linear transformation of different layers, respectively.

The encoder is composed of four successive applications of two  $3 \times 3$  convolutional layers and one downsampling operation. Each convolutional layer has a kernel size of  $3 \times 3$ , a stride of 1, and a padding of 1, and is followed by the batch normalization operation [49] and the rectified linear (ReLU) activation function [50]. The batch normalization is utilized to speed up training process and improve network performance [49]. The downsampling operation is implemented by a max-

pooling layer with a kernel size of  $2 \times 2$  and a stride of 2. The encoder with multi-scale layers extracts the hierarchical features of the input raw B-scan with different receptive fields.

The decoder includes four successive applications of one upsampling operation, a concatenation with the corresponding feature map from the encoder, and two  $3 \times 3$  convolutional layers. The upsampling operation is implemented by the bilinear interpolation with a scale factor of 2. The convolutional layers are followed by the batch normalization operation and the ReLU activation function.

Instead of directly connecting the encoder features and decoder features that aims at avoiding the loss of target response information due to the downsampling operation in the encoder [51], the RDBs are introduced between the encoder and decoder. It is found that the feature maps in the encoder, especially the low-level feature maps, include clutter information that has not been fully eliminated. Directly concatenating these encoder features to the decoder features could bring unwanted clutter information into the decoder, thus resulting in unsatisfactory clutter removal performance. Therefore, the feature maps in the encoder are first passed through a RDB to adaptively preserve the features related to the target response and reduce the unwanted clutter features, and then forwarded to the decoder.

The framework of the RDB is illustrated in Fig. 5. It includes three densely connected layers, one local feature fusion layer, and one residual learning operation [48]. In the densely connected layer, let  $f_0$  and  $f_l$  denote the input feature map and the output feature map of the  $l^{th}$  layer, respectively, the input of the  $l^{th}$  layer is the feature maps of all preceding layers  $f_0, f_1, \dots, f_{l-1}$ . The non-linear transformation of the

$l^{th}$  layer  $H_l(\cdot)$  is implemented by a  $3 \times 3$  convolutional layer followed by the ReLU. The operation in the  $l^{th}$  layer can be expressed as:

$$f_l = H_l([f_0, f_1, \dots, f_{l-1}]), \quad (1)$$

where  $[\cdot]$  is the concatenation of the feature maps along the channel dimension. Given that  $f_0$  has  $c$  channels and there are three densely connected layers, the channel number of  $f_1$ ,  $f_2$ , and  $f_3$  is set to  $c/3$ , which is referred as growth rate of RDB. The local feature fusion layer uses a  $1 \times 1$  convolutional layer to adaptively fuse the outputs of the densely connected layers with the input feature map. The operation is formulated as:

$$f_{LF} = H_{LF}([f_0, f_1, f_2, f_3]), \quad (2)$$

where  $H_{LF}$  denotes a  $1 \times 1$  convolution operation. The output channel number of  $f_{LF}$  is reduced to  $c$ . Residual learning is introduced in the last stage to combine the fused feature map with the input feature map via channel-wise addition. The final output of the RDB is obtained as:

$$f_{out} = f_0 + f_{LF}. \quad (3)$$

The RDB adaptively preserves the object reflection features while suppressing the clutter features in the feature maps from the encoder, which improves the network's capability in removing background clutter while restoring target responses.

In the final stage of the CR-Net, two  $3 \times 3$  convolutional layers and one  $1 \times 1$  convolutional layer are employed to estimate the clutter-free B-scan. The number of channels and the size of all feature maps in the network are denoted in Fig. 4.

### B. Loss Function

To train the CR-Net, a loss function combining the MAE loss and the MS-SSIM loss [52], [53] is used, which can be expressed as:

$$L = \text{MAE}(y, \hat{y}) + L_{\text{MS-SSIM}}(y, \hat{y}). \quad (4)$$

The loss function measures the dissimilarity between the predicted clutter-free image of the network  $y$  and the ground-truth image  $\hat{y}$ .

The MAE loss is calculated as:

$$\text{MAE}(y, \hat{y}) = \frac{1}{H \times W} \sum_{i,j} |y_{i,j} - \hat{y}_{i,j}|, \quad (5)$$

where  $H \times W$  is the dimension of the image, and  $i$  and  $j$  are the indices for a pixel value.

The MS-SSIM measures the perceived image quality of  $y$  compared to its ground-truth image  $\hat{y}$  in  $M$  scales [52]. The image at the scale  $k \in [1, M]$  is obtained by  $(k-1)$  iterations of a low-pass filter and a down-sampling operation with a factor of 2. The MS-SSIM is defined as:

$$\text{MS-SSIM}(y, \hat{y}) = [l_M(y, \hat{y})]^{\alpha_M} \cdot \prod_{k=1}^M [c_k(y, \hat{y})]^{\beta_k} [s_k(y, \hat{y})]^{\gamma_k}. \quad (6)$$

In Eq. (6),  $l(y, \hat{y})$ ,  $c(y, \hat{y})$ , and  $s(y, \hat{y})$  are the measures of luminance, contrast, and structure comparison given by

$$l(y, \hat{y}) = \frac{2\mu_y\mu_{\hat{y}} + C_1}{\mu_y^2 + \mu_{\hat{y}}^2 + C_1}, \quad (7)$$

$$c(y, \hat{y}) = \frac{2\sigma_y\sigma_{\hat{y}} + C_2}{\sigma_y^2 + \sigma_{\hat{y}}^2 + C_2}, \quad (8)$$

$$s(y, \hat{y}) = \frac{\sigma_y\hat{y} + C_3}{\sigma_y\sigma_{\hat{y}} + C_3}, \quad (9)$$

where  $\mu_y$  and  $\mu_{\hat{y}}$  are the means of  $y$  and  $\hat{y}$ ,  $\sigma_y$  and  $\sigma_{\hat{y}}$  are the variances of  $y$  and  $\hat{y}$ , respectively.  $\sigma_{y\hat{y}}$  is the covariance of  $y$  and  $\hat{y}$ .  $C_1 = (k_1 R)^2$ ,  $C_2 = (k_2 R)^2$ , and  $C_3 = C_2/2$  are small constants, and  $R$  is the dynamic range of the pixel values in the image. The subscript  $k$  of the luminance, contrast, and structure comparison measures in Eq. (6) denotes the assessment at the  $k$  scale. The exponents  $\alpha_M$ ,  $\beta_k$ , and  $\gamma_k$  are used to adjust the relative importance of different components at different scales. The values are calibrated and provided in [52]. The MS-SSIM loss is calculated as

$$L_{\text{MS-SSIM}}(y, \hat{y}) = 1 - \text{MS-SSIM}(y, \hat{y}). \quad (10)$$

The utilization of MAE loss yields better convergence compared to using conventionally adopted MSE loss, and the MS-SSIM loss improves the visual similarity between the predicted image and the ground truth. The combined loss function in Eq. (4) is experimentally verified to strengthen the network's clutter removal capability under the same training dataset. The comparison of network performance with different losses will be presented in Section IV.C.

## IV. EXPERIMENTS

In this section, the implementation details of the CR-Net are described in subsection A. The clutter removal performance of the CR-Net on the simulated and measured radargrams in multiple scenarios is presented in subsection B. The comparison of the clutter removal performance with existing methods is also provided. Subsection C presents the comparison of the clutter removal performance of different network architectures. Ablation study is performed in subsection D to demonstrate the effects of the inclusion of hybrid sub-datasets into the training data, the integration of the RDB in the network architecture, and the use of different losses. The generalization capability of the CR-Net in a wider range of scenarios is examined in subsection E.

### A. Implementation Details of the CR-Net

The 1,920 sets of synthetic data as mentioned in Section II.A are split into training and testing datasets in an 80%:20% ratio. The 1,536 sets of synthetic training data are augmented to 4,608 sets by selected [1, 64], [9, 72], and [13, 76] ranges of A-scans as B-scan radargrams. The augmented synthetic training data are combined with the two hybrid sub-datasets as the final training data, which includes 14,608 raw B-scans with their corresponding clutter-free ground-truth images. All the radargrams are resized to 256×64 and normalized into a range of [0, 1].

The 384 synthetic testing data, together with raw radargrams collected in different environments and radargrams in open-source GPR datasets are used to evaluate the clutter removal performance of the CR-Net, which is presented in subsection B.

The CR-Net is implemented with PyTorch on an NVIDIA 2080Ti GPU. The weights are initialized with the standard Gaussian function. The mini-batch size is set to 40. End-to-end training is performed based on the loss function defined in Eq. (4). The optimization of the network is performed using the ADAM optimizer [54] with default parameters. The initial learning rate is set to 0.0001 and decreases by a factor of 10 every 30 iterations. The model is trained for 100 epochs from scratch.

### B. Experimental Results

The performance of the well-trained CR-Net is examined on the synthetic testing data and the measured testing data. Moreover, the CR-Net is implemented to remove clutter on radargrams in open-source GPR datasets to validate its generalization capability in real-world scenes. The performance of the subspace-based method SVD [19]–[21], the LRSD method RPCA [26], and the deep-learning based CAE [33] is presented for comparison. The SVD is implemented by removing the largest singular value of the B-scan matrix. The hyperparameter  $\lambda$  in RPCA is selected as  $3 \times 10^{-2}$ . The CAE model is built and implemented following the network design in [33], and it is trained on the same training dataset as the proposed CR-Net.

1) *Results on the Simulated Testing Data:* As the simulated testing data have corresponding clutter-free images as ground truth, the clutter removal performance of different methods is evaluated using four full-reference-based metrics, including the MAE defined in Eq. (5), the MS-SSIM defined in Eq. (6), the MSE, and the peak signal-to-noise ratio (PSNR). The latter two metrics are defined as

$$\text{MSE} = \frac{1}{H \times W} \sum_{i,j} (y_{i,j} - \hat{y}_{i,j})^2, \quad (11)$$

$$\text{PSNR (dB)} = 10 \cdot \log_{10} \left( \frac{1}{\text{MSE}} \right). \quad (12)$$

In the Eqs. (11) and (12),  $y$  and  $\hat{y}$  are the normalized predicted clutter-free B-scan and the ground truth with a dimension of  $H \times W$ , and  $i$  and  $j$  are the indices for a pixel value. Smaller MAE and MSE values and larger PSNR and MS-SSIM values suggest a higher similarity between the predicted clutter-free image and the ground truth, representing a better clutter removal performance.

The comparisons of different methods using the four evaluation metrics on the simulated testing data are listed in Table III. The metric values are the averaged results of all testing data. As shown in Table III, the CR-Net outperforms other methods by a large margin under each evaluation metric. CAE and RPCA achieve the second and third best results, respectively. The conventional method SVD performs the worst.

The visual results of the CR-Net and other comparison methods on the simulated testing data are presented in Fig.

TABLE III  
QUANTITATIVE COMPARISONS OF THE PERFORMANCE OF DIFFERENT CLUTTER REMOVAL METHODS. (THE BEST RESULTS ARE HIGHLIGHTED IN BLUE.)

| Methods                            | SVD    | RPCA   | CAE    | CR-Net (ours) |
|------------------------------------|--------|--------|--------|---------------|
| MAE $\times 10^4$ ( $\downarrow$ ) | 59.62  | 47.82  | 29.62  | <b>7.59</b>   |
| MSE $\times 10^4$ ( $\downarrow$ ) | 1.79   | 1.40   | 0.38   | <b>0.04</b>   |
| PSNR ( $\uparrow$ )                | 39.82  | 41.50  | 45.77  | <b>54.92</b>  |
| MS-SSIM ( $\uparrow$ )             | 0.9515 | 0.9638 | 0.9907 | <b>0.9990</b> |

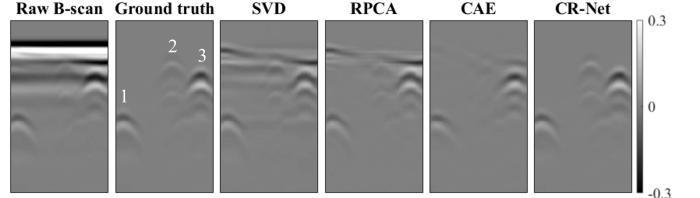


Fig. 6. Visual comparison of the clutter removal performance of different methods on the simulated testing data. The simulated scenario is three objects (1, 2, and 3 shown in ground truth) at different positions in heterogeneous soil with rough surface and water puddles. Compared with the ground truth, the CR-Net produces the best clutter-free radargram with indistinguishable clutter and intact object responses. (The radargrams are normalized into the range of [-0.5 0.5] for illustration.)

6. One example of the most challenging scenario is examined where three objects are at different positions in heterogeneous soil with a rough surface and water puddles. The marked 1, 2, and 3 in the ground truth are the reflection signatures of the three subsurface objects. Compared with the ground truth, SVD and RPCA cannot fully remove the clutter caused by the uneven surface, water puddles, and soil heterogeneity, leaving pronounced clutter that interfered with reflection signature of the shallowly buried targets 2 and 3. CAE removes most of the clutter, but it also weakens the strength of the reflected signal of object 2, which is unwanted in the clutter removal process. The CR-Net, on the other hand, not only eliminates clutter but also well restores the target responses, producing the most promising clutter-free image. The visual comparison together with the quantitative assessment in Table III demonstrates the superior performance of the CR-Net in removing clutter and restoring target responses.

2) *Results on Measured Data in Different Scenarios:* Since the ground-truth clutter-free images in the field experiments cannot be obtained, we use the improvement factor  $I_m$  to quantitatively evaluate the clutter removal performance on measured radargrams. The improvement factor is defined as the ratio of the signal-to-clutter ratio of the processed radargram  $SCR_{processed}$  to the signal-to-clutter ratio of the raw radargram  $SCR_{raw}$  [55]:

$$I_m(\text{dB}) = 20 \cdot \log_{10} \frac{SCR_{processed}}{SCR_{raw}}. \quad (13)$$

The signal-to-clutter ratio (SCR) is calculated as the ratio of the maximum amplitude of the target reflected signal  $A_{max\_signal}$  to the maximum amplitude of the clutter  $A_{max\_clutter}$  in the GPR B-scan:

$$SCR = \frac{A_{max\_signal}}{A_{max\_clutter}}. \quad (14)$$

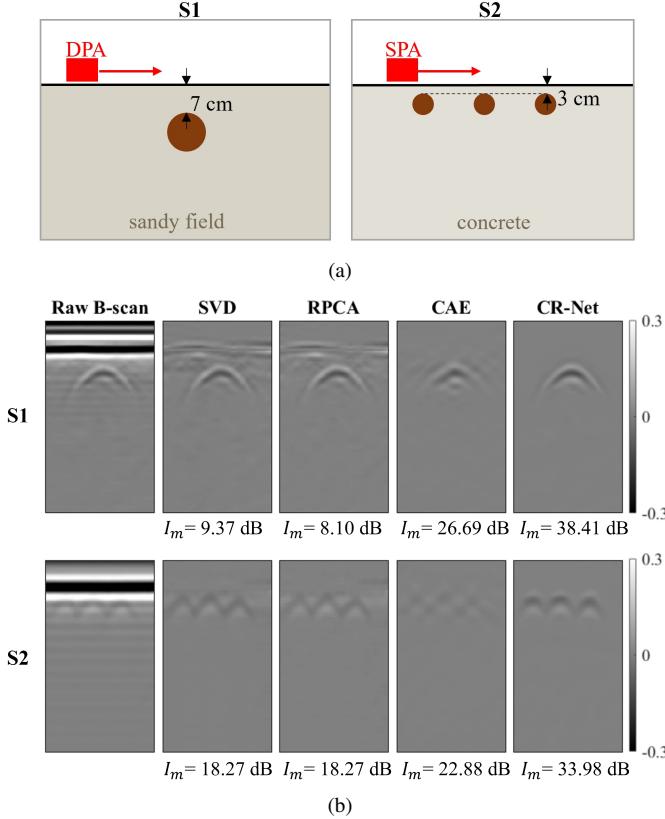


Fig. 7. (a) Illustration of the experimental scenarios S1 and S2. (b) Comparison of the clutter removal performance of different methods on the measured radargrams. The improvement factor  $I_m$  of different methods are also presented. Compared with other methods, the CR-Net successfully removes most of the clutter while restoring object reflections, and it achieves the highest  $I_m$  values in the two cases.

A higher  $I_m$  value corresponds to a better clutter removal performance.

The radargrams are collected in different scenarios using different GPR configurations, as illustrated in Fig. 7(a). In scenario 1 (S1), an aluminum can is buried in a sandy field at a depth of 7 cm. The DPA as described in Section II.B is used to collect the object reflection. In scenario 2 (S2), three rebars are located at a cover depth of 3 cm in a concrete sample, and the SVA as described in Section II.C is employed to collect the rebar reflected signals. The raw B-scans in the two cases are shown in Fig. 7(b). As the rebars in S2 are at shallow depths, a majority of their reflected signals are disguised by the strong surface clutter, which poses great challenges to extract object responses.

The clutter removal performance of different methods and their  $I_m$  scores are compared in Fig. 7(b). SVD and RPCA cannot completely suppress the background clutter, leaving residual clutter merged with object reflected signals, which is more obvious in the results of S2. CAE is effective in removing clutter, but it distorts or weakens parts of the target responses. The CR-Net, in contrast, successfully removes most of the clutter and well preserves the object responses. The high performance of CR-Net is more obvious in S2, where the interference of strong surface clutter on the reflection signatures of shallowly buried targets is significantly suppressed, and the

target reflections are well extracted. The superior performance of the CR-Net can also be quantified using the  $I_m$  values. The CR-Net achieves the highest  $I_m$  values in both cases, whereas CAE remains as the second-best method and RPCA and SVD have the lowest  $I_m$  values.

The comparison of results on the measured radargrams in this subsection shows that the CR-Net maintains its superior performance in removing real-world clutter and restoring object responses even when the object reflection is severely obscured by clutter.

3) *Results on the Open-Source GPR data:* The CR-Net is also implemented in open-source GPR radargrams to examine its generalization capability of removing clutter in completely new real-world scenarios. The raw radargrams of scenario 3 (S3) and scenario 4 (S4) are from the rebar dataset [56], and the TU1208 dataset [57], respectively. In S3, the data were collected on a concrete building using the 1500-MHz NJJ-105 GPR. As shown in the raw B-scan of S3 in Fig. 8, the rebar reflection is mixed with a horizontally interfering signal, which challenges the rebar recognition and characterization [56]. In S4, the 200-MHz GSSI system was used to survey a site filled with gneiss as depicted in Fig. 8. The site mainly consists of two dolmens, an empty concrete pipe, and three layers of pipes with an empty steel pipe, a PVC pipe full of water, and an empty PVC pipe per layer [57].

The clutter removal performance of the CR-Net and other methods for the two scenarios is compared in Fig. 8. The  $I_m$  scores of different methods are calculated and shown in the figure. In S3, SVD cannot fully remove the horizontal bandings and it affects the rebar reflected signals. This is because the clutter and rebar signals have similar strength in the raw data, so they cannot be well separated using different singular values. RPCA removes most of the clutter, but it also severely distorts the rebar reflection. CAE and CR-Net both mitigate the clutter, but CR-Net achieves better performance in restoring the rebar reflections. In S4, the clutter removal capability of SVD is the worst with pronounced residual clutter and additional horizontal bandings. RPCA and CAE suppress most of the clutter, but they result in incomplete or weakened target responses. The CR-Net achieves the best performance in removing most of the clutter while keeping the reflection signatures of subsurface objects intact. The superior performance of the CR-Net is also quantified by the highest  $I_m$  scores in both scenarios.

It is noted that the S3 and S4 are completely new scenarios, and the GPR systems employed are different from any GPR systems we have experimented with. Therefore, the clutter of S3 and S4 has different characteristics from our measured clutter. Inspiringly, the CR-Net maintains its effectiveness in removing background clutter and restoring target responses. The results shown in this subsection demonstrate the good generalization capability of the CR-Net in eliminating clutter in various real-world scenarios.

### C. Comparison of Clutter Removal Performance of Different Network Architectures

After demonstrating the superior clutter removal performance of the CR-Net over the conventional methods, its

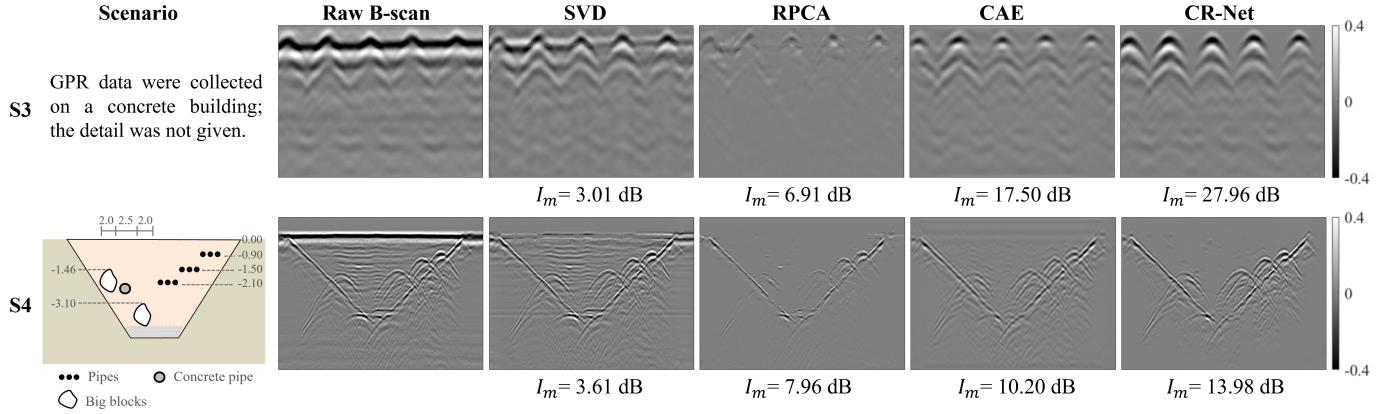


Fig. 8. Clutter removal performance of different methods on radargrams from open-source datasets. The radargram of scenario 3 (S3) and scenario 4 (S4) are from the rebar dataset [56], and the TU1208 dataset [57], respectively. In S3, the data were collected on a concrete building using the handy search NJJ-105 GPR. In S4, the 200-MHz GSSI system was used to survey a gneiss site with multiple subsurface objects. The  $I_m$  values of the results are also presented. The CR-Net achieves the best performance in eliminating clutter and restoring target responses.

TABLE IV  
COMPARISON OF CLUTTER REMOVAL PERFORMANCE OF DIFFERENT NETWORK ARCHITECTURES. (THE BEST RESULTS ARE HIGHLIGHTED IN BLUE.)

| Network architectures | HR-Net [58] | CED-Net | Base U-Net [51] | CR-Net (ours) |
|-----------------------|-------------|---------|-----------------|---------------|
| MAE $\times 10^4$ (↓) | 11.93       | 13.34   | 9.01            | <b>7.59</b>   |
| MSE $\times 10^4$ (↓) | 0.06        | 0.09    | 0.04            | <b>0.04</b>   |
| PSNR (↑)              | 52.75       | 51.98   | 54.21           | <b>54.92</b>  |
| MS-SSIM (↑)           | 0.9983      | 0.9979  | 0.9989          | <b>0.9990</b> |

performance is also compared with that of different baseline deep learning network architectures. The compared network architectures include the high-resolution network (HR-Net) [58], the convolutional encoder-decoder network (CED-Net), and the base U-Net [51]. For a fair comparison, we keep a similar amount of computational cost of each network. Specifically, the HR-Net is built using only the high-resolution convolution stream and the CED-Net is configured as the encoder-decoder structure in the base U-Net without the skip connections. The compared network architectures are trained using the same training settings, loss function, and training dataset as the proposed CR-Net. Their performance on both the synthetic testing data and the measured data is quantitatively and qualitatively compared in Table IV and Fig. 9, respectively.

It can be seen from Table IV and Fig. 9 that the base U-Net achieves better clutter removal performance than the HR-Net and the CED-Net on both the synthetic testing data and the measured data. This is because the skip connections in the U-Net preserve more complete information of the target response during the clutter removal process. Our proposed CR-Net performs the best in removing clutter and restoring target responses quantitatively and visually, as the addition of residual dense blocks adaptively preserves the features related to the target response and reduces the unwanted clutter features in the skip connection paths. The comparisons demonstrate the superiority of the base U-Net architecture and the further performance enhancement of our final CR-Net.

#### D. Ablation Study

Ablation studies are conducted to demonstrate the effectiveness of the main parts of our network design and implementation. The effects of the inclusion of hybrid sub-datasets into the training data, the integration of RDBs in the network architecture, and the utilization of different losses are investigated. The ablated models are retrained while keeping the same settings as the final model CR-Net except for the ablated parts. The performance of the ablated models on the synthetic testing data is quantitatively compared in Table V, and the performance on the measured data is visually compared in Fig. 10.

**Effect of adding hybrid sub-datasets into training data.** Model A is the network trained only on the simulated dataset. Comparing the performance of Model A and the CR-Net trained on both the simulated and hybrid datasets in Fig. 10, it can be observed that the CR-Net exhibits a significant improvement in removing clutter in real-world GPR radargrams. Using only the simulated dataset as training data cannot allow the network to learn complex and diverse distributions of real clutter, resulting in a limited clutter removal capability of Model A. This is evident in Fig. 10 where the results of Model A have obvious unremoved clutter and artifacts. The hybrid sub-datasets built with clutter data collected by different GPR systems in multiple scenarios specifically addresses this issue. Adding hybrid datasets into training data enables the network to learn the random and complex distributions of real clutter, thereby greatly enhancing the clutter removal performance of the network in real-world radargrams.

**Effect of integrating RDBs in the network architecture.** Model B is the base U-Net [51] without the integration of RDBs. As listed in Table V, the performance of Model B falls behind the CR-Net on the simulated testing data. As shown in Fig. 10, although Model B successfully removes most of the clutter, it distorts the target responses in S2 and S3. This could be due to the interference of clutter information from the skip-connected encoder features. Integrating RDBs into the skip connection paths helps to filter out the unwanted

TABLE V  
COMPARISON OF CLUTTER REMOVAL PERFORMANCE IN ABLATION STUDY (THE BEST RESULTS ARE HIGHLIGHTED IN BLUE.)

| Models                | Effect of hybrid data |   | Effect of RDB |        |        | Effects of different losses |   |   | Final model |
|-----------------------|-----------------------|---|---------------|--------|--------|-----------------------------|---|---|-------------|
|                       | A                     | B | C             | D      | E      |                             |   |   |             |
| Hybrid sub-datasets   |                       |   | ✓             |        |        | ✓                           | ✓ | ✓ | ✓           |
| RDB                   | ✓                     |   |               | ✓      |        | ✓                           | ✓ | ✓ | ✓           |
| MSE loss              |                       |   |               | ✓      |        |                             |   |   |             |
| MAE loss              | ✓                     |   | ✓             |        |        |                             |   |   | ✓           |
| MS-SSIM loss          | ✓                     | ✓ |               |        |        |                             | ✓ | ✓ |             |
| MAE $\times 10^4$ (↓) | 9.89                  |   | 9.01          | 15.93  | 10.11  | 8.66                        |   |   | 7.59        |
| MSE $\times 10^4$ (↓) | 0.08                  |   | 0.04          | 0.09   | 0.05   | 0.09                        |   |   | 0.04        |
| PSNR (↑)              | 51.37                 |   | 54.21         | 50.89  | 53.57  | 51.34                       |   |   | 54.92       |
| MS-SSIM (↑)           | 0.9985                |   | 0.9989        | 0.9973 | 0.9986 | 0.9987                      |   |   | 0.9990      |

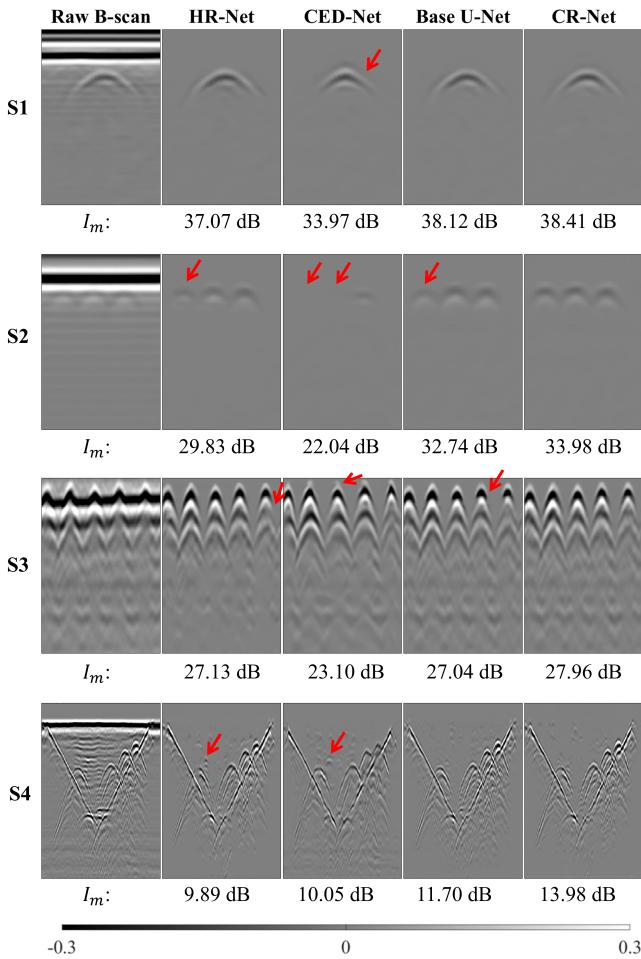


Fig. 9. Visual comparison of the clutter removal results of different network models. The improvement factor  $I_m$  values of different methods are also presented. The HR-Net [58], CAE-Net and base U-Net [51] leave residual clutter or distort parts of the target responses in some cases, as indicated by the red arrows. The final CR-Net achieves the best performance in removing clutter and restoring target responses.

clutter information while preserving target responses in the concatenated features, thereby improving the clutter removal capability of the network, as demonstrated in the results of the CR-Net shown in Fig. 10.

**Effect of using different losses.** Although the MSE loss is the dominant loss in most deep learning-based methods, it does not appear to be the best loss in the GPR clutter

removal task. This is because MSE penalizes large errors but is more tolerant of small errors [53], which could not drive the network to obtain a complete restoration of the target response. In the ablation study, the performance of the network trained using three different losses is compared. Model C-E are the network trained using the MSE loss, MAE loss, and MS-SSIM loss, respectively. As shown in Table V, Model C trained using the MSE loss has the worst score in each evaluation metric compared with Models D and E, showing the disadvantage of the MSE loss in the clutter removal task. As shown in Fig. 10, Model C either leaves residual clutter or deteriorates target responses in most cases, which does not produce favorable clutter-free results. In contrast, Models D and E produce cleaner radargram and better preserve target signals in most scenarios. This comparison result suggests that the MAE loss and MS-SSIM loss can be more suitable loss functions than the MSE loss for the clutter removal task. However, although outperform the MSE loss, they still leave slight clutter or distort target reflections in some cases. The loss function combining the MAE loss and the MS-SSIM loss leverages the advantages of the two loss functions. It drives the network to achieve consistent and superior clutter removal performance in different scenarios, as demonstrated quantitatively and qualitatively in Table V and Fig. 10.

The ablation study proves that by integrating RDBs into the network architecture, training the network using the CLT-GPR dataset, and employing the combination of MAE and MS-SSIM losses to drive network optimization, the CR-Net enjoys superior and robust clutter removal capability in diverse scenarios.

#### E. Generalization Study

After demonstrating the superior performance of the CR-Net for clutter removal in different environments using both simulated and measured data, additional investigations are conducted to verify the effectiveness of the CR-Net in a wider range of scenarios. These scenarios include the no-target case, the multi-layer case, the landmine case, and the GPR survey using different antennas, operating frequencies, and antenna-to-ground distances.

**No-target case.** The measured data in the uneven sandy field and the concrete environment without any subsurface targets as described in Sections II.B and II.C are used to examine the performance of the CR-Net in the no-target cases. The raw

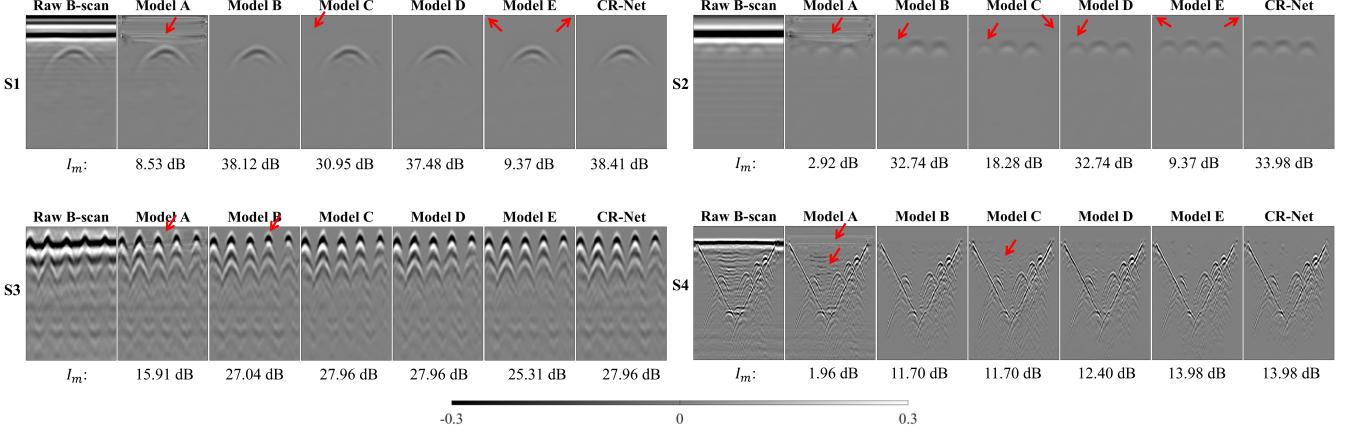


Fig. 10. Visual comparison of the clutter removal results of ablation models. The ablated models leave unremoved clutter, produce noticeable artifact, or distort parts of the target responses in some cases, as indicated by the red arrows. The final CR-Net achieves the best clutter removal capability with the clearest background and well-restored target responses. (Zoom in for a better view.)

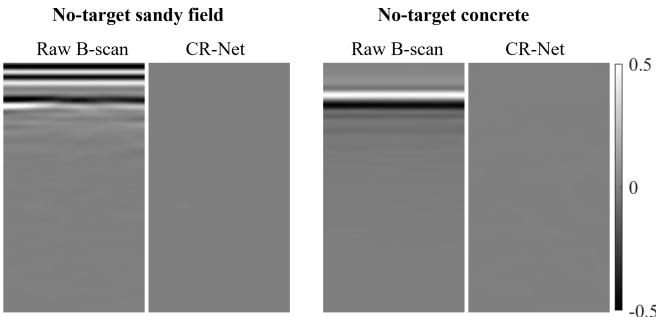


Fig. 11. The clutter removal performance of the CR-Net in no-target cases. The raw data are collected in the uneven sandy field and the concrete environment as described in Sections II.B and II.C. The CR-Net successfully removes environmental clutter without producing false targets and artifacts.

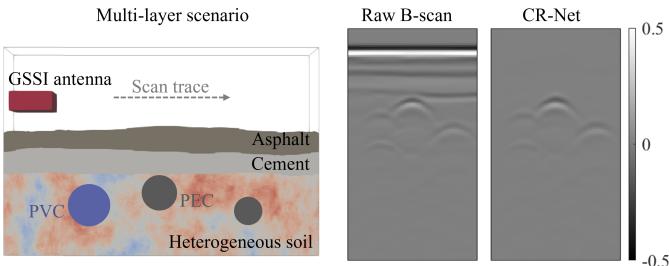


Fig. 12. The clutter removal performance of the CR-Net in the multi-layer scenario. The CR-Net removes the surface reflection at layer interfaces and well preserves the reflection of elongated targets, even when the interface between layers is rough and uneven.

radargrams and the processed radargrams by the CR-Net are shown in Fig. 11. It can be seen that the CR-Net successfully removes the ground reflection and subsurface environmental clutter without producing false targets and artifacts.

**Multi-layer case.** The performance of the CR-Net is examined in a multi-layer scenario shown in Fig. 12. The scenario consists of a 5-cm-thick asphalt layer with a relative permittivity of 5 and a conductivity of 0.001, a 5-cm-thick cement layer with a relative permittivity of 7 and a conductivity of

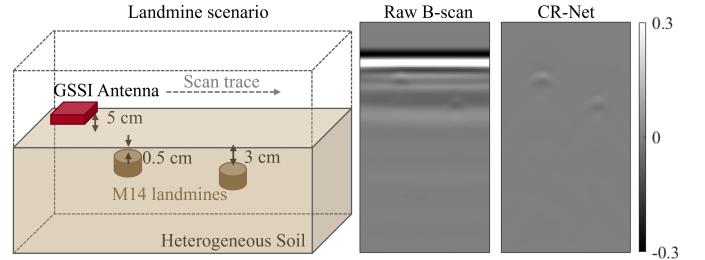


Fig. 13. The clutter removal performance of the CR-Net in the scenario with shallowly buried landmines. The CR-Net removes the environmental clutter and restores the reflection of the landmines.

0.001, and a 20-cm thick heterogeneous soil layer, which is modeled to simulate the urban road environment [41]. Each layer has a rough surface with a height variation of 4 cm. One PVC pipe and two PEC pipes are buried within the depth range of 10–25 cm. The GPR survey setup is the same as described in Section II.A. The clutter removal performance of the CR-Net in the multi-layer scenario is shown in Fig. 12. The CR-Net removes the surface reflection at layer interfaces and well preserves the reflection of elongated targets, even when the interfaces between layers are rough and uneven. The results demonstrate the applicability of the CR-Net in removing clutter and restoring target reflections in the multi-layer scenario.

**Landmine case.** The clutter removal performance of the CR-Net is examined in a scenario with shallowly buried landmines as shown in Fig. 13. Two cylindrical objects with a height of 4 cm, a diameter of 5.6 cm, a relative permittivity of 3.0 and a conductivity of 0.01 are modeled to simulate the M14 plastic-cased anti-personnel landmine [33]. They are buried in heterogeneous soil at shallow depths of 0.5 cm and 3 cm, respectively. The raw radargram and the processed radargram by the CR-Net are shown in Fig. 13. The environmental clutter is successfully removed and the reflection signatures of the landmines are well restored. The results verify that the CR-Net maintains its effectiveness in clutter removal in scenarios

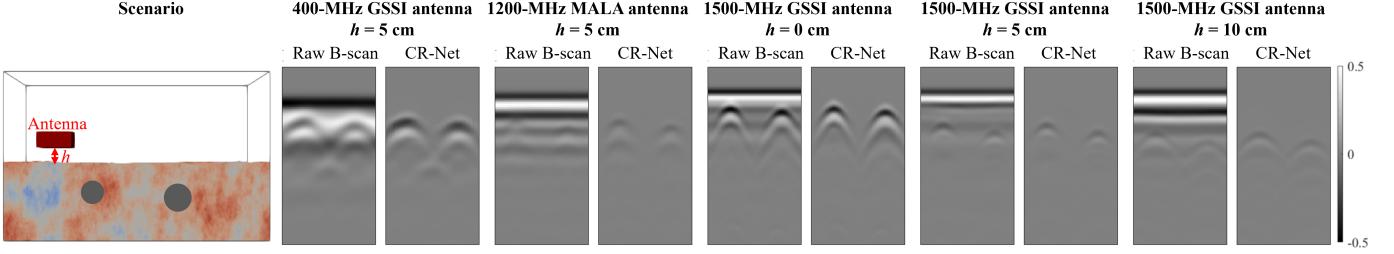


Fig. 14. The clutter removal performance of the CR-Net in the GPR survey using different antennas, operating frequencies, and antenna-to-ground distances. The CR-Net effectively and consistently removes environmental clutter in different cases.

with exposed targets such as landmines.

**GPR survey using different antennas, operating frequencies, and antenna-to-ground distances.** As GPR surveys use different antennas with different operating frequencies and have different antenna-to-ground distances in real-world applications, we further investigate the performance of the CR-Net in these cases. The subsurface scenario is the heterogeneous soil with a rough surface and two buried PEC pipes, as shown in Fig. 14. Three different antennas operating at different frequencies are used as the transmitter and receiver. They are the 400-MHz GSSI antenna [59], the 1200-MHz MALA antenna [45], and the 1500-MHz GSSI antenna [46]. For the 1500-MHz GSSI antenna, three different antenna-to-ground distances are modeled, which are 0 cm, 5 cm, and 10 cm, respectively. The raw B-scans and the processed B-scans by the CR-Net in different cases are shown in Fig. 14. Although the clutter and the reflected signatures of the targets vary with the employed antennas, operating frequencies, and antenna-to-ground distances, the CR-Net removes the clutter and restores the object reflection with consistently high performance. It is noted that the 400-MHz GSSI antenna and the 1200-MHz MALA antenna are not used in generating the training dataset, yet the CR-Net maintains its effectiveness in these cases. Therefore, the results demonstrate the versatility of the CR-Net in processing the radargrams collected by different antennas at different frequencies and with different antenna-to-ground distances.

The performance of the CR-Net in diverse scenarios illustrated in this subsection, together with the experimental results shown in Sections IV.A and IV.B, demonstrates that the performance of the CR-Net is not constraint by employed antennas, GPR operation frequencies, antenna-to-ground distances, and subsurface environments. We will examine the performance of proposed method in more diverse scenarios in our future study.

## V. CONCLUSION

In this paper, a novel clutter removal method based on the neural network is presented. A network architecture, CR-Net, is designed to eliminate clutter and restore target responses in GPR B-scans. A large-scale dataset containing diverse and complex real-world clutter is built to train the network. The clutter diversity in the dataset greatly improves the generalization capability of the network in removing clutter in

real-world radargrams. The effects of different loss functions on the network's clutter removal performance are explored, and a loss function that is more suitable for the clutter removal task is experimentally verified. Extensive experiments demonstrate that given real-world radargrams, the CR-Net can automatically and effectively remove clutter and restore target responses with consistently high performance. The proposed method outperforms the existing methods in eliminating clutter in diverse real-world scenarios and it does not require manual selection of hyperparameters.

The CR-Net with its robust clutter removal performance serves as a powerful tool to process real-world radargrams. It produces clutter-free images with well-restored target responses, which can facilitate further interpretation of radargrams such as target localization and characterization.

## REFERENCES

- [1] H. M. Jol, *Ground Penetrating Radar Theory and Applications*. Amsterdam, The Netherlands: Elsevier, 2008.
- [2] D. J. Daniels, *Ground Penetrating Radar*, 2nd ed. London, U.K.: IEE Press, 2004.
- [3] K. Iizuka, A. P. Freundorfer et al., "Step-frequency radar," *J. Appl. Phys.*, vol. 56, no. 9, pp. 2572-2583, Nov. 1984.
- [4] A. D. Olver and L. G. Cuthbert, "FMCW radar for hidden object detection," *IEE Proc. F Radar Signal Process.*, vol. 135, no. 4, pp. 354-361, Aug. 1988.
- [5] S. L. Earp, E. S. Hughes, T. J. Elkins, and R. Vickers, "Ultra-wideband ground-penetrating radar for the detection of buried metallic mines," *IEEE Aerosp. Electron. Syst. Mag.*, vol. 11, no. 9, pp. 30-39, Sep. 1996.
- [6] S. Vitebskiy, L. Carin, M. A. Ressler, and F. H. Le, "Ultra-wideband, short-pulse ground-penetrating radar: Simulation and measurement," *IEEE Trans. Geosci. Remote Sens.*, vol. 35, no. 3, pp. 762-772, May 1997.
- [7] T. Miwa, M. Sato and H. Niitsuma, "Subsurface fracture measurement with polarimetric borehole radar," *IEEE Trans. Geosci. Remote Sens.*, vol. 37, no. 2, pp. 828-837, Mar. 1999.
- [8] C.-C. Chen, S. Nag, W. D. Burnside, J. I. Halman, K. A. Shubert, and L. Peters, "A stand-off focused-beam land mine radar," *IEEE Trans. Geosci. Remote Sens.*, vol. 38, no. 1, pp. 507-514, Jan. 2000.
- [9] R. O. Schmidt, "Multiple emitter location and signal parameter estimation," *IEEE Trans. Antennas Propag.*, vol. 34, no. 3, pp. 276-280, Mar. 1986.
- [10] H. Yamada, M. Ohmiya, Y. Ogawa, and K. Itoh, "Superresolution techniques for time-domain measurements with a network analyzer," *IEEE Trans. Antennas Propag.*, vol. 39, no. 2, pp. 177-183, Feb. 1991.
- [11] E. K. Walton, "Comparison of Fourier and maximum entropy techniques for high-resolution scattering studies," *Radio Sci.*, vol. 22, no. 3, pp. 350-356, May 1987.
- [12] J. Capon, "High resolution frequency-wavenumber spectrum analysis," *Proc. IEEE*, vol. 57, no. 8, pp. 1408-1418, Aug. 1969.
- [13] R. Roy, and T. Kailath, "ESPRIT - Estimation of signal parameters via rotational invariance techniques," *IEEE Trans. Acoust. Speech. Signal Process.*, vol. 36, no. 7, pp. 984-995, Jul. 1989.

- [14] A. M. Bruckstein, T. -J. Shan, and T. Kailath, "The resolution of overlapping echoes," *IEEE Trans. Acoust. Speech. Signal Process.*, vol. 33, no. 6, pp. 1357-1367, Dec. 1985.
- [15] R. Carriere and R. Moses, "High resolution radar target modeling using a modified Prony estimator," *IEEE Trans. Antennas Propag.*, vol. 40, no. 1, pp. 13-18, January 1992.
- [16] H. Brunzell, "Detection of shallowly buried objects using impulse radar," *IEEE Trans. Geosci. Remote Sens.*, vol. 37, no. 2, pp. 875-886, Mar. 1999.
- [17] A. M. Zoubir, I. J. Chant, C. L. Brown, B. Barkat, and C. Abeynayake, "Signal processing techniques for landmine detection using impulse ground penetrating radar," *IEEE Sens. J.*, vol. 2, no. 1, pp. 41-51, 2002.
- [18] A. Benedetto, F. Tosti, L. B. Ciampoli and F. D'Amico, "An overview of ground-penetrating radar signal processing techniques for road inspections," *Signal Processing*, vol. 132, pp. 201-209, 2017.
- [19] F. Abujarad, G. Nadim, and A. Omar, "Clutter reduction and detection of landmine objects in ground penetrating radar data using singular value decomposition (SVD)," in *Proc. 3rd Int. Workshop Adv. Ground Penetrating Radar (IWAGPR)*, Delft, The Netherlands, 2005, pp. 37-42.
- [20] U. S. Khan and W. Al-Nuaimy, "Background removal from GPR data using eigenvalues," in *Proc. 13th Int. Conf. Ground Penetrat. Radar (GPR)*, Lecce, Italy, 2010, pp. 1-5.
- [21] R. Solimene, A. Cuccaro, A. Dell'Aversano, I. Catapano, and F. Soldovieri, "Ground clutter removal in GPR surveys," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 7, no. 3, pp. 792-798, Mar. 2014.
- [22] B. Karlsen, J. Larsen, H. B. D. Sorensen, and K. B. Jakobsen, "Comparison of PCA and ICA based clutter reduction in GPR systems for anti-personal landmine detection," in *Proc. 11th IEEE Signal Process. Workshop Stat. Signal Process.*, Singapore, 2001, pp. 146-149.
- [23] F. Abujarad and A. Omar, "GPR data processing using the component-separation methods PCA and ICA," in *Proc. Int. Workshop Imag. Syst. Techn.*, Minori, Italy, 2006, pp. 60-64.
- [24] H. H. Sun, Y. H. Lee, W. Luo, G. Ow, M. L. M. Yusof, and A. C. Yucel, "Dual-cross-polarized GPR measurement method for detection and orientation estimation of shallow buried elongated object," *IEEE Trans. Instrum. Meas.*, vol. 70, pp. 1-12, Sep. 2021.
- [25] D. Kumlu and I. Erer, "Clutter removal in GPR images using non-negative matrix factorization," *J. Electromagn. Waves Appl.*, vol. 32, no. 16, pp. 2055-2066, Nov. 2018.
- [26] D. Kalika, M. T. Knox, L. M. Collins, P. A. Torrione, and K. D. Morton, "Leveraging robust principal component analysis to detect buried explosive threats in handheld ground-penetrating radar data," *Proc. SPIE*, vol. 9454, p. 94541D, May 2015.
- [27] X. Song, D. Xiang, K. Zhou, and Y. Su, "Improving RPCA-based clutter suppression in GPR detection of antipersonnel mines," *IEEE Geosci. Remote Sens. Lett.*, vol. 14, no. 8, pp. 1338-1342, Aug. 2017.
- [28] X. Song, D. Xiang, K. Zhou, and Y. Su, "Fast prescreening for GPR antipersonnel mine detection via go decomposition," *IEEE Geosci. Remote Sens. Lett.*, vol. 16, no. 1, pp. 15-19, Jan. 2019.
- [29] D. Kumlu and I. Erer, "Ground penetrating radar clutter removal via randomized low rank and sparse decomposition for missing data case," *Int. J. Remote Sens.*, vol. 41, no. 19, pp. 7680-7699, Oct. 2020.
- [30] D. Kumlu and I. Erer, "Improved clutter removal in GPR by robust nonnegative matrix factorization," *IEEE Geosci. Remote Sens. Lett.*, vol. 17, no. 6, pp. 958-962, Jun. 2020.
- [31] E. Temlioglu and I. Erer, "Clutter removal in ground-penetrating radar images using morphological component analysis," *IEEE Geosci. RemoteSens. Lett.*, vol. 13, no. 12, pp. 1802-1806, Dec. 2016.
- [32] Z. K. Ni, J. Pan, C. Shi, S. Ye, and G. Fang, "DL-based clutter removal in migrated GPR data for detection of buried target," *IEEE Geosci. Remote Sens. Lett.*, early access. 2021.
- [33] E. Temlioglu, I. Erer, "A novel convolutional autoencoder-based clutter removal method for buried threat detection in ground-penetrating radar," *IEEE Trans. Geosci. Remote Sens.*, early access, 2021.
- [34] M. Moalla, H. Frigui, A. Karem and A. Bouzid, "Application of convolutional and recurrent neural networks for buried threat detection using ground penetrating radar data," *IEEE Trans. Geosci. Remote Sens.*, vol. 58, no. 10, pp. 7022-7034, Oct. 2020.
- [35] P. Bestagini, F. Lombardi, M. Lualdi, F. Picetti, and S. Tubaro, "Landmine detection using autoencoders on multipolarization GPR volumetric data," *IEEE Trans. Geosci. Remote Sens.*, vol. 59, no. 1, pp. 182-195, Jan. 2021.
- [36] 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. 2021.
- [37] H. H. Sun, Y. H. Lee, C. Li, G. Ow, M. L. M. Yusof, and A. C. Yucel, "The orientation estimation of elongated underground objects via multi-polarization aggregation and selection neural network," *IEEE Geosci. Remote Sens. Lett.*, vol. 19, Apr. 2021.
- [38] H. H. Sun, Y. H. Lee, Q. Dai, C. Li, G. Ow, M. L. M. Yusof, and A. C. Yucel, "Estimating parameters of the tree root in heterogeneous soil environments via mask-guided multi-polarimetric integration neural network," *IEEE Trans. Geosci. Remote Sens.*, early access, 2021.
- [39] Z. Wei and X. Chen, "Deep-learning schemes for full-wave nonlinear inverse scattering problems," *IEEE Trans. Geosci. Remote Sens.*, vol. 57, no. 4, pp. 1849-1860, Apr. 2019.
- [40] H. Zhou, T. Ouyang, Y. Li, J. Liu, and Q. Liu, "Linear-model-inspired neural network for electromagnetic inverse scattering," *IEEE Antennas Wireless Propag. Lett.*, vol. 19, no. 9, pp. 1536-1540, Sep. 2020.
- [41] Z. K. Ni, S. Ye, C. Shi, C. Li, and G. Fang, "Clutter suppression in GPR B-scan images using robust autoencoder," *IEEE Geosci. Remote Sens. Lett.*, early access, Oct. 2020.
- [42] H. Zhou, Y. Wang, Q. Liu, and Y. Wang, "RNMF-guided deep network for signal separation of GPR without labeled data," *IEEE Geosci. Remote Sens. Lett.*, early access, 2021.
- [43] 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, 2016.
- [44] C. Warren et al., "A CUDA-based GPU engine for gprMax: Open source FDTD electromagnetic simulation software," *Comput. Phys. Commun.*, vol. 237, pp. 208 - 218, Apr. 2019.
- [45] C. Warren and A. Giannopoulos, "Creating finite-difference time-domain models of commercial ground-penetrating radar antennas using Taguchi's optimization method", *Geophysics*, vol. 76, no. 2, pp. G37-G47, 2011.
- [46] I. Giannakis, A. Giannopoulos, and C. Warren, "Realistic FDTD GPR antenna models optimised using a novel linear/nonlinear full waveform inversion," *IEEE Trans. Geosci. Remote Sens.*, vol. 57, no. 3, pp. 1768-1778, Mar. 2019.
- [47] E. D. Guy, J. J. Daniels, S. J. Radzevicius, and M. A. Vendl, "Demonstration of using crossed dipole GPR antennae for site characterization," *Geophys. Res. Lett.*, vol. 26, no. 22, pp. 3421-3424, Nov. 1999.
- [48] Y. Zhang, Y. Tian, Y. Kong, B. Zhong, and Y. Fu, "Residual dense network for image super-resolution," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR)*, 2018, pp. 2472-2481.
- [49] S. Ioffe and C. Szegedy, "Batch normalization: Accelerating deep network training by reducing internal covariate shift," in *Proc. ICML*, Lille, France, 2015, pp. 448-456.
- [50] V. Nair and G. E. Hinton, "Rectified linear units improve restricted boltzmann machines," in *Proc. ICML*, Haifa, Israel, 2010, pp. 807-814.
- [51] O. Ronneberger, P. Fischer, and T. Brox, "U-net: Convolutional networks for biomedical image segmentation," in *Proc. Med. Image Comput. Comput.-Assisted Intervention*, 2015, pp. 234-241.
- [52] Z. Wang, E. P. Simoncelli, and A. C. Bovik, "Multi-scale structural similarity for image quality assessment," in *Proc. IEEE Asilomar Conf. Signals, Syst. Comput.*, Nov. 2003, pp. 1398-1402.
- [53] H. Zhao, O. Gallo, I. Frosio, and J. Kautz, "Loss functions for image restoration with neural networks," *IEEE Trans. Comput. Imaging*, vol. 3, no. 1, pp. 47-57, Mar. 2017.
- [54] D. P. Kingma and J. Ba, "Adam: A method for stochastic optimization", in *Proc. Int. Conf. Learn. Represent.*, 2015.
- [55] M. I. Skolnik, "MTI radar," in *Radar Handbook*, 2nd ed., New York: McGraw-Hill, 1990.
- [56] Z. Xiang, G. Ou, and A. Rashidi, "Robust cascaded frequency filters to recognize rebar in GPR data with complex signal interference," *Autom. Constr.*, vol. 124, pp. 103593, Apr. 2021.
- [57] X. Dérobert and L. Pajewski, "TU1208 open database of radargrams: The dataset of the IFSTTAR geophysical test site," *Remote Sens.*, vol. 10, no. 4, pp. 530, Mar. 2018.
- [58] K. Sun, B. Xiao, D. Liu, and J. Wang, "Deep high-resolution representation learning for human pose estimation," in *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jun. 2019, pp. 5693-5703.
- [59] S. Stadler and J. Igel, "A numerical study on using guided GPR waves along metallic cylinders in boreholes for permittivity sounding," in *Proc. 17th Int. Conf. Ground Penetrating Radar (GPR)*, Rapperswil, Switzerland, Jun. 2018, pp. 18-21.