

Robotic Inspection and Subsurface Defect Mapping Using Impact-Echo and Ground Penetrating Radar

Ejup Hoxha , Jinglun Feng , Diar Sanakov, and Jizhong Xiao , *Senior Member, IEEE*

Abstract—Concrete infrastructure often develops a variety of internal flaws that cannot be detected through visual inspection alone, and must be regularly inspected with other methods to maintain structural integrity. It has been demonstrated through previous studies that relying solely on a single non-destructive evaluation (NDE) method can be insufficient in providing a comprehensive evaluation of the structure's condition. In addition, manual NDE data collection can be labor-intensive for on-site engineers. This paper presents a robotic inspection system that uses vision-based positioning and tags NDE measurement with pose information to reveal and map subsurface defects. The system consists of three modules: 1) an Omni-directional robotic data collection platform equipped with a Realsense D435i camera for localization, an impact-echo (IE) sensor, and a ground penetrating radar (GPR), to perform automatic NDE data collection; 2) an IE data processing module that utilizes both learning-based and classical methods to interpret the IE data and reveal subsurface objects; 3) a GPR data processing module to reconstruct underground targets and create a 3D map for better visualization. Field testing demonstrates that the robotic system significantly increases the data collection speed, and the correlation of findings from both IE and GPR sensors give a comprehensive evaluation of concrete structures that will benefit the inspection and maintenance industry of civil infrastructure.

Index Terms—Robotics and automation in construction, sensor-based control, sensor fusion.

I. INTRODUCTION

NUMEROUS human-made infrastructures (bridges, tunnels, highways, etc.) in USA and around the world are reaching their life expectancy, and thus have strong needs for routine inspection and maintenance to ensure sustainability. In addition to visual inspection of surface flaws (e.g., cracks), inspectors need to detect subsurface defects (e.g., voids) using NDE instruments such as GPR and IE devices to determine the

structural integrity of concrete structures. The current practice of manual inspection requires human inspectors to first mark a grid map covering the area of interest, then push the GPR device along the grid lines and rely on the survey wheel encoder to trigger the GPR reading for post-processing, or tap the IE device at the intersection point of the grid to collect IE data. It is a time consuming and tedious work surveying a large area. In this paper, we introduce a robotic solution for automatic data collection and effective data analysis methods to interpret NDE data for mapping the subsurface defects. Robotic inspection offers the advantage of producing more consistent and cost-effective inspection results through the use of intelligent algorithms compared to manual inspection. Numerous efforts have been made to develop robotic systems for construction inspection [1], which has gained popularity in recent decades. Texas A&M University [2], [3] proposed a vision-facilitated underground pipeline mapping method to conduct GPR-based 3D reconstruction using a ground mobile robot.

There are several NDE techniques that can be used to assess the condition of concrete structures, including impact-echo [4], impact-sounding [5], ground penetrating radar [6], [7], ultrasonic [8], and visual inspection [9]. In this research, we focus on IE and GPR methods, which are particularly useful for identifying subsurface defects and deterioration areas, as well as changes in material properties in concrete structures.

IE is a technique originated from the pioneer works [4], [10], [11], [12] that utilizes vibrations to assess the condition of concrete structures. In recent years, many researchers started using neural network based machine learning methods for IE data analysis [13], [14], [15]. GPR is widely used in NDE industry for civil engineers to locate and map buried objects, measure pavement thickness and properties, and characterize subsurface features [16], [17]. Back-projection (BP) is the most commonly used algorithm for GPR imaging and subsurface reconstruction [18], [19].

There are several challenges to be resolved. First, previous learning-based IE methods do not generalize well when using different vibration sensors during inference. Second, the majority of prior research only use individual NDE technique for infrastructure inspection, and the single modal analysis may restrict data-driven decision making. The third challenge is how to determine the position and orientation of the NDE sensors accurately and in real-time, and synchronize with NDE measurements at each sampling step. In the current practice of NDE data collection, inspectors would either manually move the sensors along pre-marked grid lines or count on a high-precision

Manuscript received 19 December 2022; accepted 18 June 2023. Date of publication 29 June 2023; date of current version 7 July 2023. This letter was recommended for publication by Associate Editor A. Degani and Editor H. Moon upon evaluation of the reviewers' comments. This work was supported in part by the National Science Foundation (NSF) under Grants IIP-1915721, IIP-2112199, and in part by the U.S. Department of Transportation, Office of the Assistant Secretary for Research and Technology (USDOT/OST-R) through the INSPIRE University Transportation Center (<http://inspire-utc.mst.edu>) at the Missouri University of Science and Technology under Grant 69A3551747126. (Ejup Hoxha and Jinglun Feng contributed equally to this work.) (Corresponding author: Jizhong Xiao.)

Ejup Hoxha, Jinglun Feng, and Jizhong Xiao are with the EE Department, The City College of New York, New York, NY 10031 USA (e-mail: ehoxha@ccny.cuny.edu; jfeng1@ccny.cuny.edu; jxiao@ccny.cuny.edu).

Diar Sanakov is with the Mechanical Engineering Department, NYU, New York, NY 11201 USA (e-mail: ds6886@nyu.edu).

Digital Object Identifier 10.1109/LRA.2023.3290386

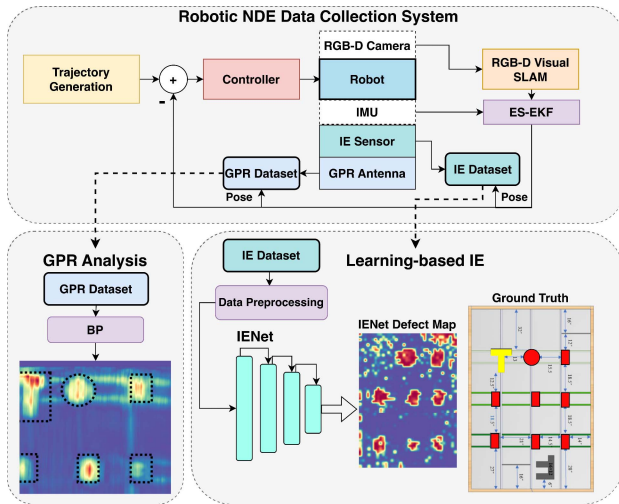


Fig. 1. System Overview: The robotic inspection system consists of three modules: a vision-based positioning and automatic NDE data collection module; an IE data processing module that generates defect maps; and a GPR data processing module that interprets the GPR data to identify subsurface objects.

global positioning system (GPS) to provide accurate position information for ground surveying in an outdoor environment.

To address these challenges, we implement a low-cost vision-based positioning method, tag the pose information on each IE/GPR sample, and develop data analysis method for IE and GPR sensors that provides a holistic solution for automated NDE data collection and subsurface defect detection and mapping.

The system overview is shown in Fig. 1, and the main contributions of this work are summarized as follows:

- An Omni-directional robot holding IE device and GPR antenna on its chassis that can move forward, backward, and sideways for fast NDE data collection.
- A multimodal sensing and NDE data analysis strategy that employs Deep Neural Networks (DNN) to interpret IE data and correlate with GPR imaging to detect and map subsurface flaws.
- A publicly accessible dataset for collected IE and GPR data over a concrete slab containing detailed ground truth information.

II. AUTOMATE NDE DATA COLLECTION VIA GROUND ROBOTS

A. Robotic Data Collection System

In this section, we introduce our robotic inspection system that facilitate and automate the NDE data collection. Fig. 2 shows a rendered model of the robot that uses Mecanum wheels [20] to generate a holonomic system, making the robot easy to move sideways and survey the ground along grid patterns while maintaining its orientation. By preserving a constant orientation, the robot enhances the resilience of the visual SLAM. A front-facing Realsense RGB-D camera with embedded IMU is utilized for vision-based positioning and localization. A sliding mechanism is installed at the rear of the robot that host IE device assembly in the sliding plate. The dome in Fig. 2 is used to amplify impact-sounding (IS) signals and insulate ambient noise. The IS

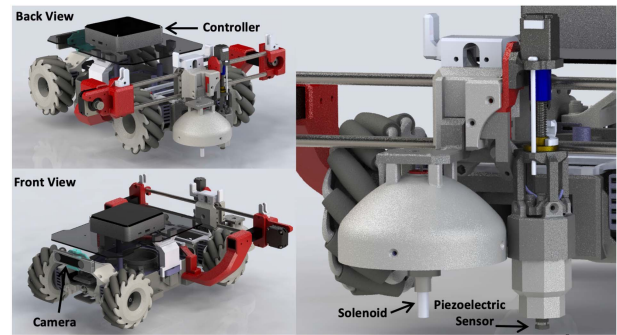


Fig. 2. Omni-directional robot for automatic IE data collection.

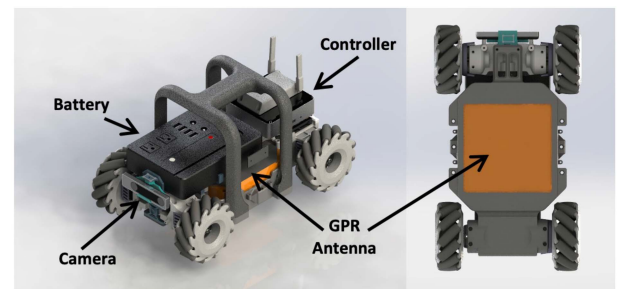


Fig. 3. Omni-directional robot for automatic GPR data collection, where a GPR antenna is installed at the bottom of the robot chassis.

acoustic signal is collected by a microphone which is installed at the focal point of the dome to detect shallow defects. An electric solenoid is installed at the center of the dome to tap the ground. Next to the dome, an up-and-down damping mechanism is installed to host a piezoelectric (PZT) sensor for collecting vibration signal when the PZT sensor is in contact with ground surface. The damping mechanism uses a ball-screw shaft to move the PZT sensor along vertical axis and prevent it from wear and tear when it is lift off the ground. The entire IE/IS assembly is mounted on a sliding plate to facilitate the data collection at multiple locations along the guide rail.

Additionally, as shown in Fig. 3, we develop an identical robot equipped with a PavScan GPR antenna manufactured by Geophysical Survey System Inc. (GSSI). The GPR antenna is installed at the bottom of robot's chassis with 2 cm space clearance above the ground. The sole distinction is that this GPR robot does not show the sliding track for mounting IE/IS assembly. Instead, it is installed with a rechargeable battery and a high-level controller (Intel NUC) to power the device and synchronize the pose data with GPR scan data.

B. Vision-Based Positioning

In this section, we present the vision-based positioning approach employed by the robotic system. An Intel D435i RGB-D camera with integrated IMU is added to the robotic system, enabling accurate and robust pose estimations. To further increase pose output frequency and robustness required for GPR data collection, we utilize an Error-State Extended Kalman Filter (ES-EKF) [21]. We fuse the SLAM pose obtained from the

camera during the update step with the IMU's pose during the prediction step of ES-EKF.

State variables include position p , orientation q , velocity v , acceleration bias a_b , gyro bias ω_b , and gravity vector g . To circumvent the Gimbal Lock issue, we represent orientation using quaternions, $q = [q_x \ q_y \ q_z \ q_w]$. The state of our system is illustrated in (1).

The Error-State Extended Kalman Filter comprises three primary states: the true state x_t , the nominal state x , and the error state δx . The true state embodies the corrected values of the states, while the nominal state updates the state according to the model in (3), or in a more straightforward form as shown in (2). Here, x_k can represent either the true state or the nominal state, and w denotes the IMU's white noise. In the nominal state, we do not account for the noises, biases, and gravity and, thus, consider them constant.

$$x = \begin{bmatrix} p & v & q & a_b & \omega_b & g \end{bmatrix} \quad (1)$$

$$\hat{x}_{k+1} = f_k(x_k, u_k, w) \quad (2)$$

$$\begin{bmatrix} p_{k+1} = p_k + v_k \cdot \Delta t + \frac{1}{2} \Delta t^2 \cdot (R(a_m - a_b) + g) \\ v_{k+1} = v_k + \Delta t \cdot (R(a_m - a_b) + g) \\ q_{k+1} = q_k \otimes q\{(\omega_m - \omega_b)\Delta t\} \\ a_b = a_b \\ \omega_b = \omega_b \\ g = g \end{bmatrix} \quad (3)$$

Upon receiving measurements from the IMU, we update the states according to the nominal state in (3) and the covariance according to (4). We denote our nominal state at the current step as \hat{x}_k . Operator \otimes represents quaternion multiplication defined as $q_1 \otimes q_2 = [q_1]_L \cdot q_2$, where $[\cdot]_L$ represents left quaternion product matrix.

$$\hat{\Sigma}_{k+1} = F \cdot \Sigma_k \cdot F^T + G \cdot Q \cdot G^T \quad (4)$$

Here, F represents the linearized state transition matrix of the nominal state, while G encompasses state propagation noise and Q is the IMU covariance matrix, which includes acceleration noise and angular velocity noise.

While the IMU update occurs during the *prediction* phase of the EKF, the camera serves as the measurement *update*. We utilize RGB-D ORB-SLAM3 [22] for visual SLAM pose generation. To perform the measurement update, we first calculate the Kalman Gain using (5). Once we obtain the Kalman Gain, we need to compute the error state using (7), where Z_{k+1} represents the measurement from visual SLAM and x_k denotes the current nominal or true state, depending on whether we had a measurement correction in the previous step or not. The error-state kinematics are shown in (6), where V_i , A_i , and W_i are white Gaussian noises typically provided in the sensor datasheet, $[\cdot]_x$ denotes skew operator.

$$K_{k+1} = \hat{\Sigma}_{k+1} H_{k+1}^T (H_{k+1} \hat{\Sigma}_{k+1} H_{k+1}^T + R_{k+1})^{-1} \quad (5)$$

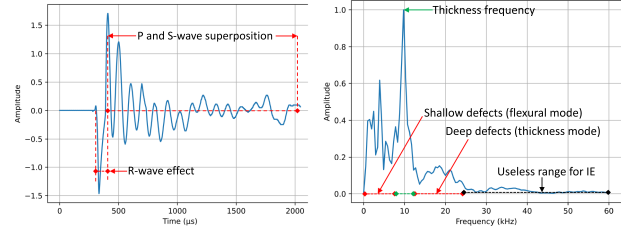


Fig. 4. IE signal collected over experimental concrete slab of CCNY Robotics Lab. Time domain signal on the left composed of P, R, and S-wave, and its frequency domain transfer on the right with approximate important regions. We can clearly see highest peak around 9.7 kHz which corresponds to the thickness frequency.

$$\begin{bmatrix} \delta p_{k+1} = \delta p_k + \delta v_k \cdot \Delta t \\ \delta v_{k+1} = \delta v_k + \Delta t \cdot (-R[a_m - a_b]_x \cdot \delta \theta + g) + V_i \\ \delta \theta_{k+1} = R^T\{(\omega_m - \omega_b)\Delta t\} \delta \theta - \delta \omega_b \Delta t + \Theta_i \\ \delta a_b = \delta a_b + A_i \\ \delta \omega_b = \delta \omega_b + W_i \\ \delta g = \delta g \end{bmatrix} \quad (6)$$

After determining the error state, we inject the observed state into the nominal state, resulting in the true state as shown in (8).

$$\delta x_{k+1} = K_{k+1} (Z_{k+1} - H_{k+1} \hat{x}_{k+1}) \quad (7)$$

$$x_{k+1} = \hat{x}_{k+1} + \delta x_{k+1} \quad (8)$$

After correcting the nominal state, we also need to update our covariance matrix using (9).

$$\Sigma_{k+1} = \hat{\Sigma}_{k+1} - K_{k+1} H_{k+1} \hat{\Sigma}_{k+1} \quad (9)$$

In this case, the H matrix represents the camera's measurement update model.

Finally, visual SLAM provides the pose for each NDE measurement. We further implement a time synchronizer software in the on-board computer under ROS (Robot Operation System), which takes in messages of different types from multiple sources, and outputs them only if it has received a message on each of those sources with the same timestamp. It is used to synchronize the NDE measurements with vision-based positioning data.

III. CONCRETE INSPECTION AND DEFECT MAPPING

A. Impact-Echo Principle

In impact-echo testing, a mechanical impact is applied to the surface of a concrete structure, causing the structure to vibrate. The resulting vibration signal, which is composed of P-waves, S-waves, and R-waves, contains valuable information about the structure's geometry and condition. Of these waves, the P-wave is the most important for impact-echo analysis, while the R-wave is particularly useful for determining the depth of vertical cracks. The effect of the R-wave can be seen as the first and deepest dip in the impact-echo time-domain signal as shown in Fig. 4. R effect is faded fast, P and S-waves contribute to the rest of the signal. The speed of a wave in a solid object is directly proportional to the mass density of the object ρ , Young's modulus of elasticity E , and Poisson's ratio ν . The speed of a P-wave can be calculated

using (10).

$$C_P = \sqrt{\frac{E(1-v)}{\rho(1+v)(1-2v)}} \quad (10)$$

The properties of the impact will heavily affect the results of IE. A long contact time t_c with the surface will act as a low-pass filter, and filter high frequencies. Therefore, useful information about the structure, and/or the flaw will be lost. When manually collecting IE data, impact is also generated manually. Usually, a steel sphere is used. Diameter of the sphere D is directly related with the contact time t_c . On the other hand, t_c defines maximum frequency f_{\max} that will be captured by the measurement. The relationship between t_c and D is described by Hertz theory of elastic impact [23], and is given in (11).

Beside contact time and sphere diameter, the force F is another parameter that defines the impact. However, the effect of different applied impact force does not heavily affect the normalized frequency response [12].

$$t_c = \frac{0.0043D}{h^{0.1}} \rightarrow t_c \approx 0.0043D \quad (11)$$

The relationship between contact time t_c , and maximum frequency f_{\max} , is defined experimentally (12). However, it could also be derived by using force-time function [12].

$$f_{\max} = \frac{1.25}{t_c} \rightarrow f_{\max} = \frac{1.25}{0.0043D} \quad (12)$$

IE method uses principle of wave reflection. When a wave changes medium, or encounters discontinuities will reflect. The reflected wave will contain information about the depth of the discontinuity. A direct relationship between P-wave velocity C_p , thickness frequency $f_{thickness}$, and depth d of the defect/structure is shown on (13). Parameter β corrects effects of S-wave, and is usually taken as $\beta = 0.96$ for concrete. This is also called as thickness mode of IE.

$$f_{thickness} = \frac{\beta C_P}{2d} \quad (13)$$

On the other hand, when the defects in concrete are shallow, we need to use a short contact time t_c in order to measure their depth. However, the maximum useful frequency f_{\max} puts a limit on the length of this contact time, as shown in (12). As a result, we cannot use this method to measure the depth of shallow defects. We call this mode of analysis the flexural vibration mode. In this mode peak frequency of IE signal is lower than the thickness frequency of the slab. To evaluate the effectiveness of this method, we generate maps showing the peak frequencies of shallow defects, which are represented as lower frequencies.

B. Learning Based Impact-Echo

In this section, we present a new model for classifying IE signals called IENet. This model outperforms existing IE learning-based methods in terms of generalization and produces impressive results on different datasets. Our focus is solely on classifying IE signals as either defective or solid concrete, rather than estimating defect depth using IE properties.

TABLE I
IENET MODEL STRUCTURE

Layer	Configurations	Output Shape
Input	IE Signal	(860 × 1)
residual_block_1	(kernel=1x200, filters=8, stride=1)	(860 × 1 × 8)
activation_1	ReLU	(860 × 1 × 8)
maxpool1d_1	(pool_size=2, stride=2)	(430 × 1 × 8)
residual_block_2	(kernel=1x100, filters=16, stride=1)	(430 × 1 × 16)
activation_2	ReLU	(430 × 1 × 16)
maxpool1d_2	(pool_size=2, stride=2)	(215 × 1 × 16)
residual_block_3	(kernel=1x50, filters=16, stride=1)	(215 × 1 × 16)
activation_3	ReLU	(215 × 1 × 16)
maxpool1d_3	(pool_size=2, stride=2)	(107 × 1 × 16)
residual_block_4	(kernel=1x25, filters=32, stride=1)	(107 × 1 × 32)
activation_4	ReLU	(107 × 1 × 32)
maxpool1d_4	(pool_size=2, stride=2)	(53 × 1 × 32)
residual_block_5	(kernel=1x13, filters=64, stride=1)	(53 × 1 × 64)
activation_5	ReLU	(53 × 1 × 64)
maxpool1d_5	(pool_size=2, stride=2)	(26 × 1 × 64)
residual_block_6	(kernel=1x7, filters=64, stride=1)	(26 × 1 × 64)
activation_6	ReLU	(26 × 1 × 64)
maxpool1d_6	(pool_size=2, stride=2)	(13 × 1 × 64)
reshape_1	(13 · 64 × 1)	(832 × 1)
bilstm_1	BiLSTM(units=32)	(64 × 1)
bilstm_2	BiLSTM(units=32)	(64 × 1)
bilstm_3	BiLSTM(units=32)	(64 × 1)
Output	Dense(2, Softmax)	2 × 1
No# of parameters	534498	

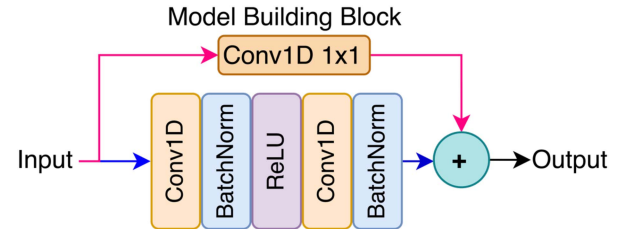


Fig. 5. Residual block used as building block in IENet.

Structure of IENet is shown in Table I. Base building block of this model is called residual block, shown in Fig. 5. We employ a residual connection strategy similar to ResNet [24], which helps the input signal to propagate faster through the network. The model consists of six residual blocks, all of which have the same architecture, but differ in terms of filters and kernel size. After the signal has passed through all residual blocks it is reshaped into a (832,1) vector. The resulting features are passed through three bidirectional long short-term memory (Bi-LSTM) layers until they reach the final dense layer and softmax activation. The number of filters and kernel size for each residual block is shown in Table I. The overall model is small, with a total of 534498 parameters. We use categorical cross-entropy for loss function, shown in (14), where p represents the predictions and t represents the targets. In this paper, we also implement and evaluate several other models that demonstrate good results. In Section IV, we explain why we ultimately choose IENet as the winning candidate.

$$L_i = - \sum_j t_{i,j} \cdot \log p_{i,j} \quad (14)$$

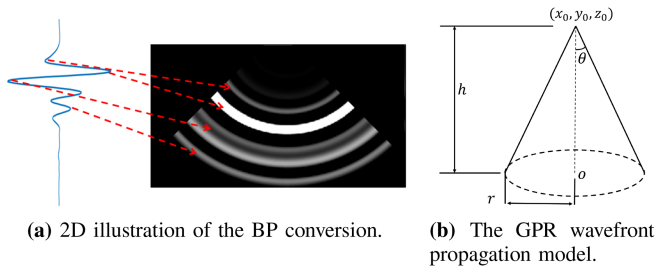


Fig. 6. An illustration of the BP principle. (a) The 2D visualization of the BP algorithm, where the left blue signal indicates an A-scan, the red-dotted lines indicate that A-scan raw data are converted into several cones; (b) The GPR wavefront transmission model, which is a shape of the cone.

C. GPR Inspection Analysis

GPR inspection is a wave propagation technique that transmits a pulse of polarized high-frequency electromagnetic (EM) waves into the subsurface media. EM wave attenuates as it travels in media and reflects when it encounters a material change. GPR antenna would thus record the strength and traveled time of each reflected pulse [17]. The reflections are then amplified, processed, and displayed on a screen as an A-scan signal, analogous to a waveform in an oscilloscope. When the GPR device moves along a straight line perpendicular to utility pipes, the ensemble of the A-scans forms a B-scan, which is shown as the hyperbolic feature, indicating the objects' locations as well [25], [26].

In practice, the EM wavefront transmission has a *circular cone* pattern as shown in Fig. 6, where the radius is r , axis length is h and the specific apex is at (x_0, y_0, z_0) , its axis is parallel to the z -axis. As illustrated in Fig. 6, the brighter area indicates the higher amplitude region in B-scan. Meanwhile, the depth of the bright area in the BP image indicates the depth of the A-scan pulse. By implementing the BP algorithm in A-scan data, the back-projected model could be represented in (15), which explicitly indicates the GPR field of view.

$$\forall \mathcal{A}_q^k \in \mathcal{B}_k, \quad (x - x_0)^2 + (y - y_0)^2 = m^2 * (v * t)^2 \quad (15)$$

We denote m as the slope of the cone. If the apex angle is 2θ , then $m = \tan(\theta)$, where θ is the semi vertical angle. So the radius of the circle r at a certain height ($h = v * t$) from the top will be $m * h$. Note that x_0, y_0, z_0 represents the location of each origin A-scan measurement. $\mathcal{A}_q^k = \{\mathbf{a}_t | t = 1, \dots, n_q\}$ represents the q -th A-scan measurement in k -th B-scan data, while scalar t and vector \mathbf{a}_t indicate the traveling time and amplitude of A-scan signal respectively, then n_q means the total samples in a A-scan measurements. Meanwhile, we also have $\mathcal{B}_k = \{\mathcal{A}_q^k | q = 1, \dots, n_k\}$ that represents the k -th B-scan consisting of n_k A-scans.

Implementing the BP algorithm on B-scan data results in transforming each A-scan data into a circular cone in 3D space, where the circular cone contains subsurface information such as potential objects or defects. However, it is essential to preprocess the raw GPR data to remove noise and enhance the signal in order to obtain clear information on the subsurface defect. In particular, we used the noise removal technique proposed in our



Fig. 7. Ground truth of concrete slab DS2. The left side shows the design of the slab while the right side shows the real slab view before concrete pouring.

previous work [27]. Then, each A-scan data can be transformed into a circular cone in 3D space through the BP process. Once combining multiple circular cones in the 3D space, a 3D volume map can be created where the intersection of the circular cones highlights potential subsurface targets or defects.

GPR BP images usually indicate the presence of objects or defects. However, the specific defect category cannot be directly determined from the GPR BP images, and identification of the defects usually relies on experts' knowledge and the background information of the survey site. For instance, circular or elliptical anomalies could represent voids or sinkholes, while linear features might indicate fractures or rebars. Complex patterns could signify structural deformation or areas with multiple defects.

To improve the accuracy of the interpretation, comparing the observed features in the GPR image with known subsurface features from other sources, such as Impact Echo feature maps, can further assist in confirming the presence of defects and rule out false positives. More detailed experimental studies regarding the identification of the defects via GPR and IE will be introduced and discussed in the next Section.

IV. EXPERIMENTAL STUDY

In this work we use GPR and IE sensors to collect data over our concrete slab. Concrete slab contains buried objects, which emulate different defects that usually happen on concrete structures. Defects are buried on different depths. In this section we discuss methodology and results in detail.

A. Datasets and IE Data Preprocessing

Datasets Introduction: In this paper, we use two different datasets, Dataset1 (DS1) and Dataset2 (DS2). Specifically, DS1 is an online public IE dataset [13] of a standard concrete slab with eight buried rectangular objects made of various materials. DS2 is a dataset we create ourselves, containing GPR and IE data collected from an artificially constructed concrete slab with defects. The defects are simulated by placing various objects, such as T-shaped metal plates, foam, and boxes over rebars at the bottom of the slab. The details of DS2 are depicted in Fig. 7.

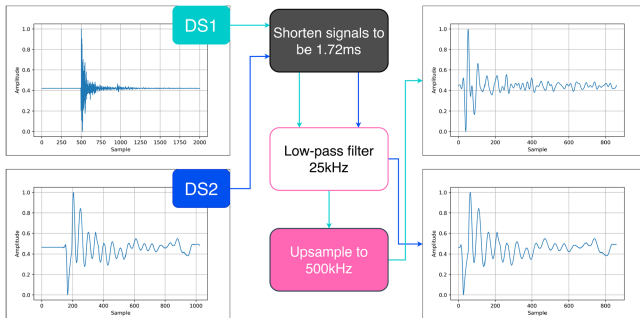


Fig. 8. Data preprocessing procedure for DS1 and DS2. Both datasets are cut, new length is 1.72 ms. Then, all data go through low-pass filter. For DS2 that is the end of the process. DS1 goes through upsampling block to match the sampling rate of DS2. Both datasets are collected over concrete slabs that are 8 in deep. Thus, both signals look very similar after the processing.

DS1 contains 2016 IE data points which are annotated from [28]. We use this dataset for training and evaluating various learning-based IE defect mapping methods. Furthermore, each IE signal on this dataset is collected over eight identical concrete slabs using a sample rate of 200 kHz, and $n = 2000$ samples per signal. Each slab contains four different types of defects, installed on different depths. Additional information about this dataset can be found in [13], [29].

We use two Omni-directional robots to collect IE and GPR data (DS2). Each IE and GPR data point is tagged with location information. In particular, there are 1824 impact-echo signals collected in total and further separated into two parts. First part has a higher resolution (31x38). For second part, we lowered the resolution from 38 IE samples per line to 34 (19x34). In addition, there are 47997 A-scan signals collected as GPR data to further generate GPR defect maps.

IE Data Preprocessing: In III-A, we described some important parameters for IE data collection. To generate the impact we use an electrically controlled impactor, which is calibrated to have a contact time of $t_c \approx 77.4 \mu\text{s}$, and allows useful frequencies up to $f_{\text{max}} \approx 16 \text{ kHz}$ for thickness mode of IE. Another important parameter is sampling frequency. We used a sampling frequency of 500 kHz, and $n = 1024$ samples per signal.

One major problem for IE data is that many researchers and NDE engineers use different parameters when collecting vibration signals. In our case, DS1 and DS2 have no common parameters in the way signals are digitized. IE signals from DS1 are 10 ms long, while signals from DS2 are 2.048 ms long. Both signals are sampled at different sample rates 200 kHz and 500 kHz, thus digitally they are vectors of different lengths. To address this issue, we first employ a preprocessing strategy that aligns the two datasets on signal duration. Specifically, we truncate the signals in DS2 to keep only $n = 860$ samples (1.72 ms) per signal, and we cut away useless parts of the signals. In DS1, we keep $n = 344$ samples per signal (1.72 ms). We then apply a low-pass filter with a cutoff frequency of 25 kHz, which is mostly effective for DS1. Finally, we upsample DS1 to match the sampling frequency of DS2. The flow diagram of this process is shown in Fig. 8.

TABLE II
COMPARISON BETWEEN DIFFERENT MODELS IMPLEMENTED ON THIS PAPER

Model	Acc. %	Defect Acc.%	Solid Acc.%	DS2 Map
BL-1DCNN	88.65	70.0	95.0	Fail
BL-BiLSTM	80.03	61.0	87.0	Fail
1DCNN	82.93	75.0	85.64	Fail
Dp-1DCNN	84.13	84.38	84.04	Weak
BiLSTM	85.71	60.93	94.15	Weak
CRNN	88.09	68.75	94.68	Good
IENet1024	87.69	60.93	96.08	Fail
IENet	90.48	75.0	95.74	Best

Quantitative results are generated using test data from DS1. “DS2 MAP” column states how well model identifies defects on DS2, shown in fig. 9. BI means baseline model from [13].

B. Training

We implemented fairly small models, hence we do not need to use expensive equipment for training. We train models on a server with Intel Core i7-7700 CPU @ 3.60 GHz CPU, Ge-Force GTX 1070 GPU, and 32 GB RAM. For training and validation we use the data collected over *Slab1* to *Slab7* from DS1. The train/validation data split rate is 72%/28%. Across all our models we used a learning rate of 0.0001, batch size of 16, and trained for 200 epochs. We employed early stopping strategy, and we only save models that perform best on evaluation data. To emphasize the importance of data preprocessing, all our models except IENet1024 are trained on preprocessed data. For training IENet1024 we truncate IE signals from DS1, and only keep $n = 1024$ most significant samples for each signal. To normalize the data we use min-max normalization.

C. Results and Evaluation

IE models comparison: In this study, we use Slab8 data from DS1 for evaluating all models in Table II. Two baseline models: Baseline 1DCNN, and Baseline BiLSTM are selected from a previous study by Dorafshan et al. [13] for comparison. It is worth noting that, as previously shown by Dorafshan et al., the baseline models are not effective in detecting defects when the data are collected by different IE sensor, i.e., the model fails to generalize.

In Table II, we show quantitative performance of all models. Defect accuracy is calculated by $\frac{TP}{TP+FN}$, where true positive TP represents number of defects when they are detected correctly, whereas false negative FN number of defects that are erroneously classified as solid concrete. Solid accuracy is defined by $\frac{TN}{TN+FP}$, where true negative TN represents number of solid concrete input that are classified correctly, whereas false positive FP represents number of solid concrete input are erroneously classified as defective. Overall accuracy is calculated by $\frac{TP+TN}{TP+FN+TN+FP}$.

Deeper 1DCNN has a bigger model depth than Baseline 1DCNN, and is the best performing model on defect classification (84.38%). On the other hand, this model does not perform as good as baseline on solid concrete classification (84.04%). Our first model that performs comparable with baseline is CRNN. This model has almost the same overall accuracy, and a slightly lower defect accuracy than baseline. IENet1024 is the only

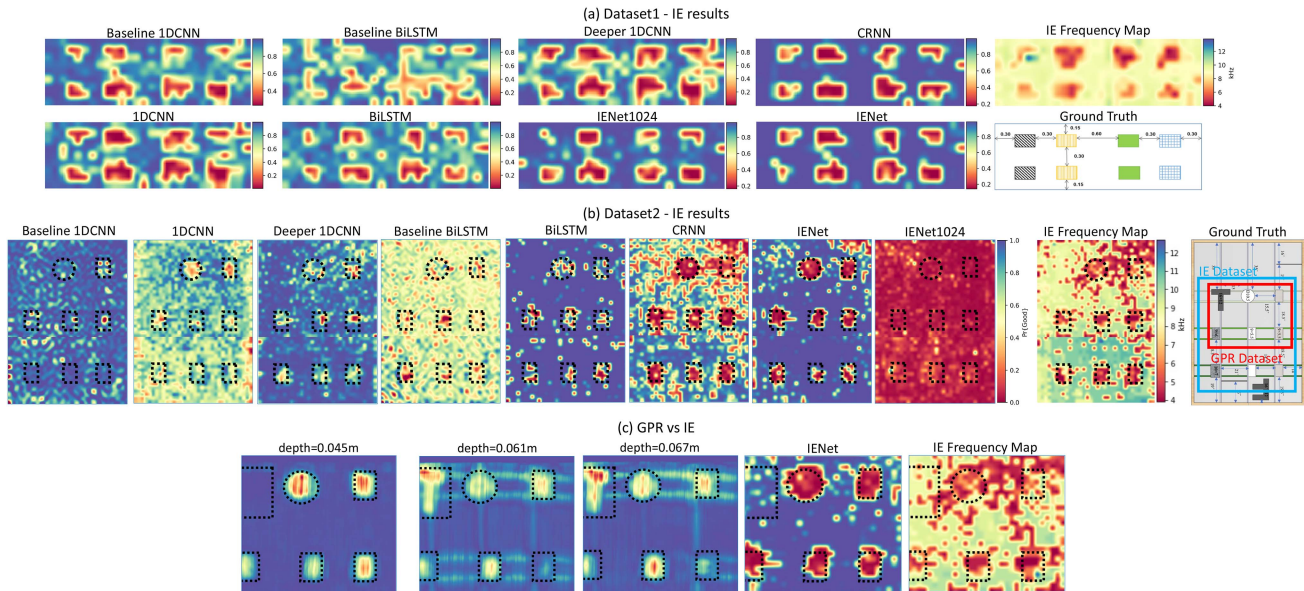


Fig. 9. Impact-echo and GPR defect mapping: (a) Test results for IE models using DS1 testing data, classical IE frequency analysis, and ground truth schematics; (b) Test results for DS2, classical IE frequency analysis, and ground truth schematics; (c) GPR results at three different depths compared to IENet and IE frequency analysis. IE data are collected over the blue rectangle area shown in ground truth schematics for DS2, while GPR covered the red rectangle area. IENet generates accurate maps for both datasets. GPR is able to detect all defects, including rebars, and T-shaped metal. Rebars and T-shaped metal are not detected by IE.

model trained using raw DS1 data as explained in Section IV-B. This model performs very good on solid concrete classification (96.08%), but not as good on defect detection (60.93%). Finally, IENet is identified as the best model which has the highest overall accuracy (90.48%), second best accuracy on detecting defects (75%), and second best accuracy on detecting solid concrete (95.74%), performing better than baseline in all categories.

To further analyze the results we generate probability maps for DS1 data for each model, shown in Fig. 9(a). When compared on the quality of the classification we can see that CRNN and IENet models are highly confident when classifying IE signals. This is a highly desirable property. On the other hand, Baseline 1DCNN shows a good performance, but is less confident on classifying IE signals. Finally, IE frequency map is able to reveal most of the defects, but requires extensive data analysis to be able to generate the map correctly.

We evaluate models on DS2, to see if they are able to generalize. For this analysis we solely rely on qualitative results, but that are clearly visible. In Fig. 9(b), we show probability maps generated from using DS2. Notice that baseline models completely fail to recognize defects. Additionally, most of our models fail too. The only model that accurately reveals all eight defects inserted in the concrete slab is IENet. The use of residual layers in the IENet model contributed to its improved performance, they help the network learn more effectively by addressing the issue of vanishing gradients.

The IENet1024 model achieved a high level of accuracy on the DS1 test data, because testing and training data are collected using one sensor. IENet and IENet1024 have almost identical model architecture. However, IENet1024 has a poor performance in DS2, classifying all concrete as defective (in red), in contrast IENet identifies all defects correctly. This suggests that

data preprocessing plays a crucial role on generalization, along the model architecture.

IE vs GPR Comparison: Finally, we compare defect maps from GPR, classical IE, and learning-based IE method. For GPR data, as described in Section III-C, we use BP-based migration method to generate the GPR imaging result, and show the 2D GPR imaging slices at different depth. GPR imaging shows amplitude differences reflected by dielectric changes. Hence, GPR is good at mapping objects or defects that have significant dielectric property changes in concrete. Fig. 9(c) shows that different objects/defects are recovered at certain depths. When comparing GPR and IE results, we see that they are correlated. However, GPR successfully reveals rebars and the T-shaped metal sheet, while IE does not. The reason is that the metal sheet buried in concrete does not have a big enough difference in vibration impedance. Moreover, rebars' radius is too small to be detected by IE. Additionally, GPR image cannot differentiate between objects and defects, while learning-based IE classifies concrete conditions being defective or solid. Classical IE reveals some of the defects, but it does not meet the bar set by learning-based IE and GPR since the IE measurement data is difficult to interpret without expert knowledge. The correlation of findings from both IE and GPR sensors give the most comprehensive evaluation of concrete structures.

V. CONCLUSION

In this study, we first introduce a robotic inspection system carrying IE and GPR sensors to automate NDE data collection. Instead of relying on a survey wheel encoder along pre-marked grid lines to trigger GPR sampling, we use visual SLAM technology to estimate the pose in real time, and tag the GPR samples

with pose information for GPR imaging using BP algorithm. This vision-based positioning and time-triggered approach enable the robot to survey the ground along arbitrary trajectory hence significantly speed up the data collection process.

Second, we study several learning-based models to classify the IE data and evaluate concrete conditions into being defective or solid. We demonstrate that IENet is the best model that outperform others and has the highest overall classification accuracy. It exhibits generalization ability in defecting subsurface defects even when the model is trained on a relatively small dataset obtained by different IE sensor at different site.

Finally, we show that BP algorithm can generate detailed 3D GPR image of concrete structures revealing defects and objects. GPR imaging can map dielectric changes in concrete structure, but it requires further study to classify the defects and objects. On the other hand, IE is able to calculate defect depths only in thickness mode but does not work in shallow mode where IE cannot determine accurate depth information. By using the proposed learning-based model for IE data analysis and BP algorithm for GPR imaging to generate a defect map, it enables inspectors to quickly assess the condition of the concrete structures. The combination of these two complementary methods is highly valuable and represents a substantial advancement in the field of subsurface defect recovery using GPR and IE. Experimental results demonstrate the effectiveness of the multi-modal sensing and data analysis strategy in subsurface defect mapping. The dataset for IE and GPR tagged with location and ground truth information will be made open source to benefit the NDE industry.

REFERENCES

- [1] H. M. La, N. Gucunski, K. Dana, and S.-H. Kee, "Development of an autonomous bridge deck inspection robotic system," *J. Field Robot.*, vol. 34, no. 8, pp. 1489–1504, 2017.
- [2] H. Li, C. Chou, L. Fan, B. Li, D. Wang, and D. Song, "Toward automatic subsurface pipeline mapping by fusing a ground-penetrating radar and a camera," *IEEE Trans. Automat. Sci. Eng.*, vol. 17, no. 2, pp. 722–734, Apr. 2020.
- [3] C. Chou, S.-H. Yeh, and D. Song, "Mirror-assisted calibration of a multi-modal sensing array with a ground penetrating radar and a camera," in *Proc. IEEE/RSJ Int. Conf. Intell. Robots Syst.*, 2017, pp. 1457–1463.
- [4] N. H. Nicholas Carino and M. Sansalone, "A point source-point receiver, pulse-echo technique for flaw detection in concrete," *ACI J. Proc.*, vol. 83, pp. 199–208, 1986.
- [5] E. Hoxha, J. Feng, D. Sanakov, A. Gjinofci, and J. Xiao, "Robotic inspection and characterization of subsurface defects on concrete structures using impact sounding," *Struct. Health Monit.*, 2021, pp. 875–884, 2021.
- [6] J. Feng, L. Yang, E. Hoxha, B. Jiang, and J. Xiao, "Robotic inspection of underground utilities for construction survey using a ground penetrating radar," *J. Comput. Civil Eng.*, vol. 37, no. 1, 2023, Art. no. 04022049.
- [7] J. Feng, L. Yang, E. Hoxha, D. Sanakov, S. Sotnikov, and J. Xiao, "Gpr-based model reconstruction system for underground utilities using gprnet," in *Proc. IEEE Int. Conf. Robot. Automat.*, 2021, pp. 845–851.
- [8] J. Krautkrämer and H. Krautkrämer, *Ultrasonic Testing of Materials*, 4th ed. Berlin, Germany: Springer, 1990.
- [9] L. Yang et al., "Deep neural network based visual inspection with 3D metric measurement of concrete defects using wall-climbing robot," in *Proc. IEEE/RSJ Int. Conf. Intell. Robots Syst.*, 2019, pp. 2849–2854.
- [10] M. Sansalone and N. J. Carino, *Impact-echo: A Method for Flaw Detection in Concrete Using Transient Stress Waves*. Washington, D.C. USA: U.S. Dept. Commerce, 1986.
- [11] M. J. Sansalone and W. B. Streett, *Impact-Echo: Nondestructive Evaluation of Concrete and Masonry*. Washington, D.C. USA: U.S. Dept. Commerce, 1997.
- [12] N. Carino, "The impact-echo method: An overview," in *Proc. World Struct. Eng. Congr.*, 2001. [Online]. Available: https://tsapps.nist.gov/publication/get_pdf.cfm?pub_id=860355
- [13] S. Dorafshan and H. Azari, "Deep learning models for bridge deck evaluation using impact echo," *Construction Building Mater.*, vol. 263, 2020, Art. no. 120109.
- [14] J. Xu, J. Zhang, and Z. Shen, "Recognition method of internal concrete structure defects based on 1D-CNN," *J. Intell. Fuzzy Syst.*, vol. 42, no. 6, pp. 5215–5226, 2022. [Online]. Available: <https://doi.org/10.3233/JIFS-211784>
- [15] F. Elghaish et al., "Developing a new deep learning CNN model to detect and classify highway cracks," *J. Eng., Des. Technol.*, vol. 20, pp. 993–1014, 2022.
- [16] S. Demirci, H. Cetinkaya, E. Yigit, C. Ozdemir, and A. A. Vertiy, "A study on millimeter-wave imaging of concealed objects: Application using back-projection algorithm," *Prog. In Electromagnetics Res.*, vol. 128, pp. 457–477, 2012.
- [17] S. Demirci, E. Yigit, I. H. Eskidemir, and C. Ozdemir, "Ground penetrating radar imaging of water leaks from buried pipes based on back-projection method," *NDT E. Int.*, vol. 47, pp. 35–42, 2012.
- [18] X. Xie, J. Zhai, and B. Zhou, "Back-fill grouting quality evaluation of the shield tunnel using ground penetrating radar with bi-frequency back projection method," *Automat. Construction*, vol. 121, Art. no. 103435.
- [19] H. Zhang, O. Shan, G. Wang, J. Li, S. Wu, and F. Zhang, "Back-projection algorithm based on self-correlation for ground-penetrating radar imaging," *J. Appl. Remote Sens.*, vol. 9, no. 1, 2015, Art. no. 095059.
- [20] N. Tlale and M. de Villiers, "Kinematics and dynamics modelling of a mecanum wheeled mobile platform," in *Proc. 15th Int. Conf. Mechatron. Mach. Vis. Pract.*, 2008, pp. 657–662.
- [21] J. Solá, "Quaternion kinematics for the error-state Kalman filter," *CoRR*, vol. abs/1711.02508, 2017.
- [22] C. Campos, R. Elvira, J. J. G. Rodríguez, J. M. Montiel, and J. D. Tardós, "ORB-SLAM3: An accurate open-source library for visual, visual-inertial, and multimap SLAM," *IEEE Trans. Robot.*, vol. 37, no. 6, pp. 1874–1890, Dec. 2021.
- [23] W. Goldsmith and J. T. Frasier, "Impact: The theory and physical behaviour of colliding solids." *J. Appl. Mechanics*, vol. 28, no. 4, pp. 639–639, Dec. 1961.
- [24] K. He, X. Zhang, S. Ren, and J. Sun, "Deep residual learning for image recognition," 2015. [Online]. Available: <http://arxiv.org/abs/1512.03385>
- [25] S. Li, H. Cai, D. M. Abraham, and P. Mao, "Estimating features of underground utilities: Hybrid GPR/GPS approach," *J. Comput. Civil Eng.*, vol. 30, no. 1, 2016, Art. no. 04014108.
- [26] C. Yuan, S. Li, H. Cai, and V. R. Kamat, "GPR signature detection and decomposition for mapping buried utilities with complex spatial configuration," *J. Comput. Civil Eng.*, vol. 32, no. 4, 2018, Art. no. 04018026.
- [27] J. Feng, L. Yang, E. Hoxha, and J. Xiao, "Improving 3D metric GPR imaging using automated data collection and learning-based processing," *IEEE Sensors J.*, vol. 23, no. 5, pp. 4527–4539, Mar. 2023.
- [28] S. Dorafshan and H. Azari, "Annotated impact echo dataset (bare decks)," *Mendeley Data*, vol. 1, 2020, doi: [10.17632/44rb96872r.1](https://doi.org/10.17632/44rb96872r.1).
- [29] D. Sattar and A. Hoda, "Evaluation of bridge decks with overlays using impact echo, a deep learning approach," *Automat. Construction*, vol. 113, 2020, Art. no. 103133.