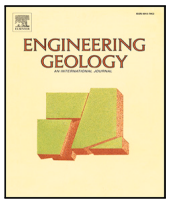
Engineering Geology 363 (2026) 108545

Contents lists available at ScienceDirect

Engineering Geology

journal homepage: www.elsevier.com/locate/enggeo



Short communication

Elastodynamic imaging of voids in a PML-truncated layered solid using a deep convolutional neural network

Boyoung Kim [a], Shashwat Maharjan [a], Bruno Guidio [d], Jacob Thomas [a], Fazle Mahdi Pranto [a], Chanseok Jeong [a,b,c,*]

[a] School of Engineering and Technology, Central Michigan University, Mount Pleasant, MI 48859, USA
[b] Earth and Ecosystem Science Program, Central Michigan University, Mount Pleasant, MI 48859, USA
[c] Institute for Great Lakes Research, Central Michigan University, Mount Pleasant, MI 48859, USA
[d] Department of Civil and Environmental Engineering, The Catholic University of America, Washington, DC 20064, USA

ARTICLE INFO

Keywords:
Void detection
Elastodynamic imaging
Element-wise classification
Deep convolutional neural network
Machine learning
Non-convolutional CFS-PML

ABSTRACT

Voids in the subsurface pose significant challenges to infrastructure stability and safety, often leading to structural failures and costly repairs. In this study, we propose a deep convolutional neural network (DCNN) framework for elastodynamic imaging of voids in a semi-infinite soil domain truncated by non-convolutional second-order complex-frequency-shifted perfectly matched layers (CFS-PML). The method employs element-wise classification to map void probabilities within the domain using elastodynamic surface responses from non-scanning type probing. The training datasets are generated using a level-set wave solver, producing input-layer features from measured surface responses and output-layer features as element grid maps indicating void probabilities. The DCNN is trained to predict the void probability in each element and reconstructs targeted voids by clustering high-probability elements. Numerical results demonstrate that, via rigorous out-of-distribution tests, the proposed DCNN can effectively detect and image voids, including those with complex shapes that were not even included in training data. The model's performance remains stable under receiver uncertainties, including measurement noise and random tilting, with noise-trained models showing notably improved robustness. It also yields reasonable predictions under sparse receiver layouts and maintains stable performance across material-property variations. Compared with full-waveform inversion, our DCNN offers more accurate reconstructions, making void locations clearer. This study highlights the potential of integrating advanced deep-learning techniques with wave propagation models for improved subsurface exploration and characterization.

1. Introduction

Detecting subsurface voids—e.g., underground cavities, karst formations, natural caves within ground structures, and voids between concrete pavement and a sublayer—is critical for maintaining the integrity and safety of various infrastructures, including tunnels, roads, and buildings (De Giorgi and Leucci, 2014). These voids can cause ground subsidence, structural failures, and catastrophic collapses. Traditional methods, such as standard penetration testing (SPT) and cone penetration testing (CPT), provide direct measurements but require invasive drilling, rendering them impractical for large-scale surveys. Non-destructive testing (NDT) techniques have emerged as a practical solution to address these challenges, enabling the detection of subsurface voids without excavation. NDT methods are particularly effective for surveying extensive areas while minimizing environmental and structural disruption (Gupta et al., 2022; Rahimi et al., 2024).

Among these NDT techniques, electrical resistivity tomography (ERT) is utilized to image underground anomalies by measuring subsurface resistivity (Saad et al., 2012; Park et al., 2014; Bharti et al., 2016). However, ERT faces limitations, particularly in environments with low resistivity contracts or high noise sensitivity, which makes it difficult to obtain high-resolution images of subsurface voids (Oldenborger et al., 2005). Ground penetrating radar (GPR) has been widely used to detect subsurface voids by utilizing electromagnetic waves (Thitimakorn et al., 2016; Luo and Lai, 2020; Dai et al., 2022). A GPR sensor device that integrates both a transmitter and receiver is used in a scanning mode over a surface. At each scan position, the device excites and receives electromagnetic waves, and the collected A-scan signals are combined to form B-scan images along each scanning line. The B-scan images provide cross-sectional views where the signal response to a target is typically strongest when the probe is positioned directly above

* Corresponding author at: School of Engineering and Technology, Central Michigan University, Mount Pleasant, MI 48859, USA.
E-mail address: jeong1c@cmich.edu (C. Jeong).

https://doi.org/10.1016/j.enggeo.2026.108545
Received 25 December 2024; Received in revised form 3 January 2026; Accepted 6 January 2026
Available online 12 January 2026

the feature. Recently, ML-based feature extraction methods have been studied for characterizing anomalies, such as a void or water-bearing cavity, from B-scan images of GPR applied at subsurface (Li et al., 2025a,b; Zhang et al., 2025a). An ML-trained feature-extraction model automatically learns characteristic patterns in the B-scan associated with specific targets and effectively enables automated defect detection. However, the method may need a second opinion for the cases when the wave signals from a lower void, out of vertically-aligned voids, is not noticeable in B-scan image or when scanning-like probing is not available due to a logistical challenge or constraints at inspected sites.

Several studies have investigated surface-based seismic waves (e.g., Rayleigh or Love waves) for imaging near-surface to detect defects such as cavities (Nasseri-Moghaddam et al., 2007; Rahnema et al., 2022), water leaks of pipes (Dashwood et al., 2020), and damage to structural cracks (Xu et al., 2022). While these surface waves can effectively probe the near-surface layers, they suffer from limited resolution and accuracy at deeper depths (Giammarinaro et al., 2023). Recognizing the need for improved subsurface characterization, advanced methods like full-waveform inversion (FWI) have been explored. FWI utilizes both surface and body waves (i.e., P and S waves), without scanning-type sensing, to provide a detailed characterization of subsurface material properties (Tran et al., 2013). FWI iteratively updates material profiles by solving an optimization problem that minimizes the misfit between observed and simulated waveforms, offering higher resolution and accuracy for detecting anomalies in the wave speed distribution (Virieux and Operto, 2009; Kallivokas et al., 2013; Kucukcoban et al., 2019; Mirzanejad et al., 2021; Liu et al., 2023; Kim et al., 2025). However, because FWI typically requires repeated forward and adjoint simulations throughout the inversion process, it leads to significant computational cost and is therefore unsuitable for real-time characterization. To overcome these limitations, Aaker et al. (2021) introduced a significant reduction in computational time by up to 92% using time-space tiling on graphics processing units (GPUs). However, despite these advancements, it remains ineffective for real-time detection. The FWI of wave velocities in the subsurface can only reconstruct a targeted void as areas of low wave speeds because the continuum solid model in the material inversion cannot accommodate the vanished wave speeds. It hardly delineates the void because the gradient of a misfit function with respect to the void's boundary information is hard to calculate analytically or semi-analytically. In efforts to address this limitation, recent advancements in machine learning (ML) have demonstrated strong potential for reducing the computational costs of traditional optimization-based approaches of FWI. For instance, Wu and Lin (2018) developed a data-driven model that utilizes CNN to perform accurate and rapid characterization of subsurface velocity distribution in a truncated domain, effectively identifying flat and curved layered profiles.

Building on these advancements, we study a data-informed ML framework to clearly image voids within a semi-infinite soil domain using a computationally-efficient and robust ML approach without using scanning-type sensing. To achieve this, first, a perfectly matched layer (PML) is employed as a wave-absorbing boundary condition, effectively truncating the unbounded physical domain and focusing computations on the region of interest (Berenger, 1994; Francois et al., 2021; Kang and Kallivokas, 2010; Kucukcoban and Kallivokas, 2011; Fathi et al., 2015; Kim and Kang, 2019, 2024; Zhang et al., 2024, 2025b). In particular, the presented work utilizes non-convolutional second-order CFS-PML (Francois et al., 2021) because of its ease of implementation, but the method is scalable to other PML formulations or absorbing boundary conditions.

Secondly, this work leverages an element-wise classification (Pranto et al., 2023; Kim et al., 2024), which has been recently developed for producing the probability map for each element being void from elastodynamic wave responses at sensors, enabling the accurate imaging of target voids with arbitrary shapes and quantities. The element-wise classification approach can address the limitations of the existing ML-based void imaging methods. Namely, ML has been applied to image internal defects by utilizing surface measurement data in thin plates (Jiang et al., 2022), laminated composites (de Assis and Gomes, 2021), and rectangular cuboids in 3D polymer structures (Niu and Srivastava, 2022). However, these methods rely on using specific shape parameters (e.g., circle diameter, ellipse size and orientation, square edge length, or V-notch angle) as output-layer features of ML models, which require prior assumptions about the shapes or the number of defects.

Overall, this study presents a DCNN architecture capable of detecting voids of arbitrary location, size, and shape in PML-truncated layered solids. The presented method requires a-prior information of background stratigraphy that can be obtained from SASW, MASW, or FWI material inversion analysis. Once the strata's material profile is determined, the presented machine learning process can be utilized to clearly image hidden voids within the strata. The presented method is gradient-free and benefits from training data produced by a level-set wave solver (Chatzi et al., 2011; Jung et al., 2013; Sun et al., 2014), which effectively models voids with vanishing wave speed.

The remaining sections of the paper are organized as follows. In Section 2, we briefly describe the level-set wave solver, which is used for forward modeling to generate training data. Then, in Section 3, we introduce the configurations of the domain that includes voids of various sizes, shapes, and locations to generate training data using the element-wise classification method. The training data consists of input-layer features (i.e., elastodynamic surface responses) and output-layer features (i.e., the element-wise classification map). Section 4 presents a methodology of a DCNN architecture and evaluation metrics. In Section 5, we present the performance of our CNN using in-distribution test datasets produced by our level-set wave solver. Finally, Section 6 presents parametric studies on out-of-distribution test datasets produced by ANSYS to evaluate the model under receiver uncertainties, sparse layouts, and material variations and provides a comparative study with FWI.

2. Forward modeling

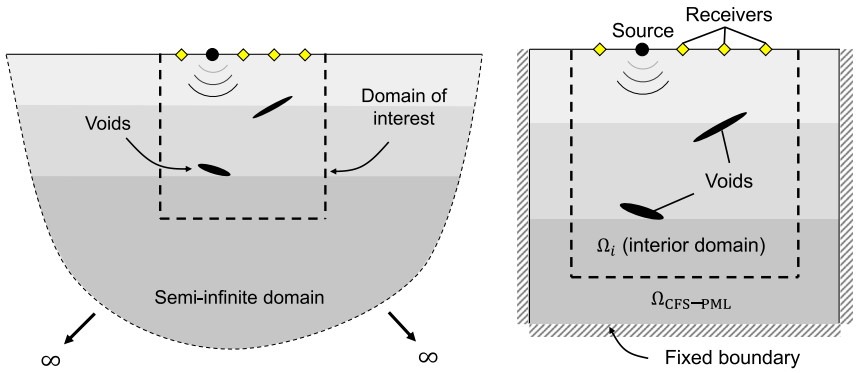
2.1. Wave physics

In this paper, we present a data-informed DCNN framework to detect underground voids within a 2D layered solid domain using wave responses, induced by a single wave source at a fixed location, measured at surficial receivers. To obtain surface responses, we simulate the propagation of elastic waves in a semi-infinite layered solid with internal voids, as shown in Fig. 1(a). We utilize the wave-absorbing boundary condition, non-convolutional second-order CFS-PML, to truncate the unbounded domain (Francois et al., 2021; Guidio et al., 2024). The semi-infinite domain has been rendered finite by introducing a wave-absorbing buffer $\Omega_{\text{CFS-PML}}$, as depicted in Fig. 1(b). Within the interior domain Ω_i, displacement responses are calculated without reflections from the truncated boundary. We assume that the domain is composed of linear, isotropic, and elastic solids with voids under plane-strain conditions. The fixed boundary conditions constrain the right, left, and bottom perimeters of the PML, and the domain's initial state is at rest.

2.2. Level-set wave solver

We apply the conventional level-set method (Jung et al., 2013; Sun et al., 2014), using the finite element approximation to compute elastodynamic surface waves in a solid containing voids as input-layer features in training data. This method does not require re-meshing when the shape of the voids is updated (Pranto et al., 2023). Namely, the solid's displacement $\mathbf{u}$ in an element is approximated as:

$$\mathbf{u}^h(x,y,t) = V(x,y)\sum_{i=1}^{N_n}\phi_i(x,y)\mathbf{u}_i(t), \tag{1}$$



(a) Semi-infinite domain. (b) PML-truncated domain.

Fig. 1. The semi-infinite domain and the PML-truncated finite computational domain.

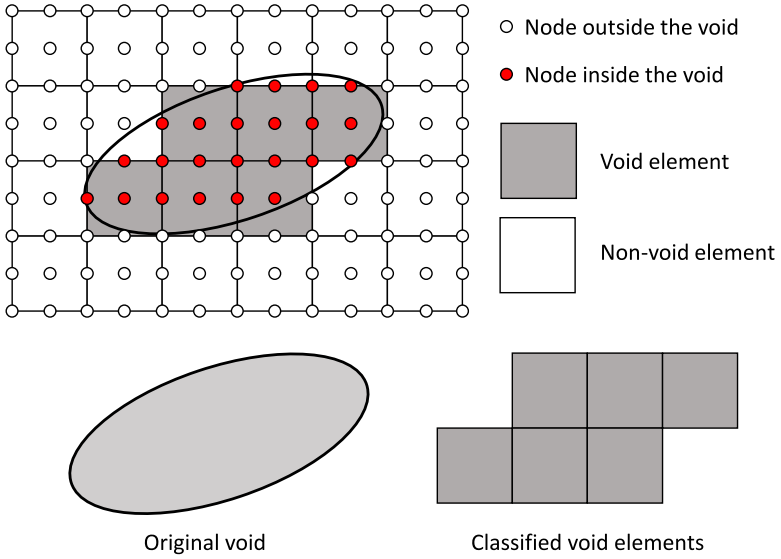


Fig. 2. An illustration of the level-set categorization of void and non-void elements for a 9-node element.

where

$$V(x,y)=\begin{cases}0, & \text{if } (x,y) \text{ is determined as a part of a void element,}\\ 1, & \text{if } (x,y) \text{ is determined as a part of a non-void element.}\end{cases} \tag{2}$$

We formulate Eq. (1) using 9-node quadrilateral elements, nodal displacements $\mathbf{u}_i$, and local shape functions ϕ_i at the ith node, where N_n denotes the number of nodes in an element. As shown in Fig. 2, elements whose centroid and at least four of their nine nodes fall inside the original void outline are marked as void elements. In this case, $V(x,y)=0$. In all other cases, an element is considered as a non-void element, $V(x,y)=1$.

Following the global assembly of the system matrices and the force vector, the governing equation for the time-dependent displacement field vector is obtained. The global solution at each time step is then computed using the Newmark integration scheme (Newmark, 1959).

2.3. Verification of our level-set wave solver

To verify our level-set wave solver, we compare its numerical solutions at receiver locations with those obtained using the third-party simulator, ANSYS Mechanical, in an enlarged domain with fixed boundaries. As shown in Fig. 3, the domain is 400×200 m, ensuring that reflections from the fixed boundaries do not enter the computational domain of interest during the simulation. We compare the displacement-time histories for domains with a circular void of its radius 5 m and its centroid at the depth of 20 m. The reference solutions are obtained using ANSYS's automatic unstructured mesh generation, which explicitly models void boundaries with 8-node quadrilateral elements of an average size of 0.5 m. The grid is arbitrarily generated in consideration of the boundary of the void, as illustrated in Fig. 4. The shear wave velocity (V_s) and dilatational wave velocity (V_p) of each layer are depicted in Fig. 8. Ricker-pulse signal $F(t)$, which is applied as a point node force signal in Eq. (3), is used by the wave source at (10, 50) m. It is utilized in a setting where its positive value is directed inward or toward the domain. The signal's maximum amplitude is 100 N/m, with $f=20$ Hz as its central frequency, as shown in Fig. 5, and the time step is 0.001 s.

$$F(t)=\begin{cases}\dfrac{100\left[\left(0.25\left(2\pi f t-3\sqrt{6}\right)^2\right)-0.5\right]e^{-0.25\left(2\pi f t-3\sqrt{6}\right)^2}-13.e^{-13.5}}{0.5+13.e^{-13.5}}, & \text{if } t\le\dfrac{6\sqrt{6}}{2\pi f},\\ 0, & \text{if } t>\dfrac{6\sqrt{6}}{2\pi f}.\end{cases} \tag{3}$$

To further examine the wave propagation behavior simulated by our level-set wave solver, Fig. 6 presents sequential displacement amplitude ($|\mathbf{u}|$) at different time instants. The wavefields radiate from the

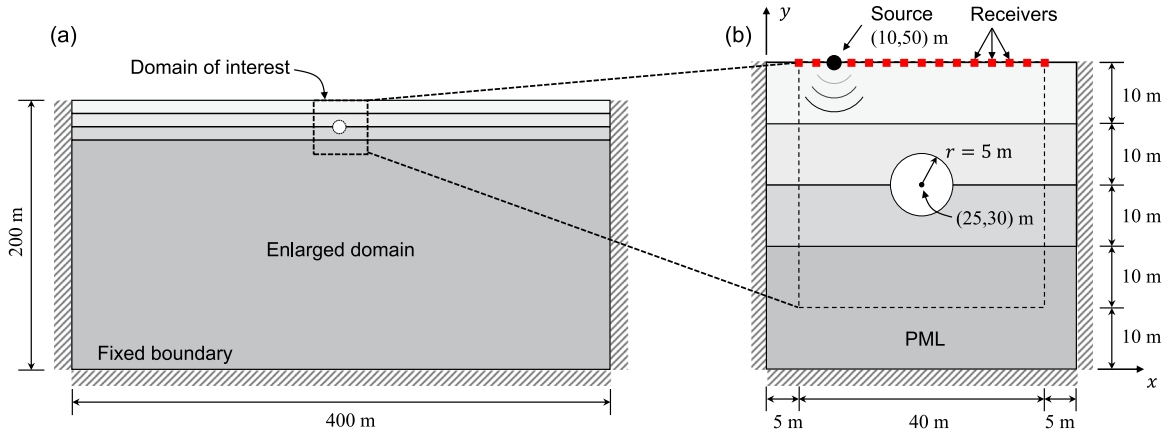


Fig. 3. (a) Enlarged domain and (b) PML-truncated domain with voids of radius 5 m.

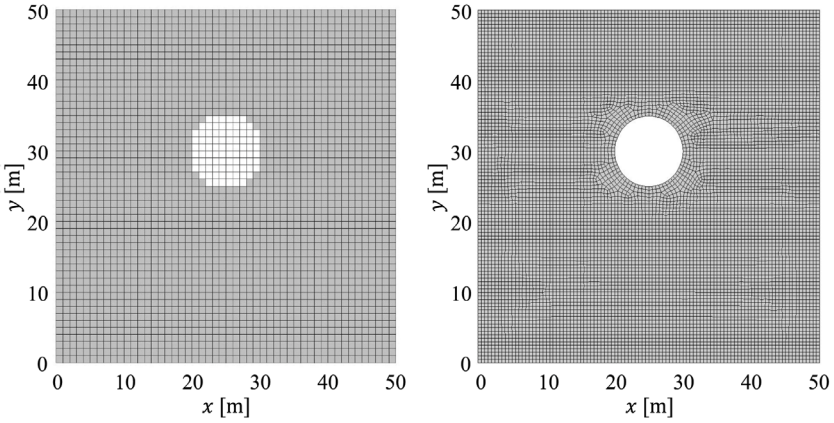


Fig. 4. (a) Structured mesh for our PML-level-set wave solver (white elements represent the void) and (b) unstructured mesh for ANSYS.

point source and propagate through the layered medium, scattering around the void boundary before being gradually absorbed by the PML without noticeable reflections from the outer boundary. These contour plots demonstrate that the proposed solver accurately reproduces wave propagation, scattering, and absorption phenomena within the computational domain. Fig. 7 presents the excellent agreement between displacement-time histories obtained using our PML-level-set solver and ANSYS at the receiver locations. Our observations indicate that the waves are effectively absorbed by the PML, and no significant reflections are observed from the fixed outer boundary. In addition, the level-set modeling is as effective as the explicit boundary meshing for modeling voids.

3. Data generation

3.1. Forward modeling parameters

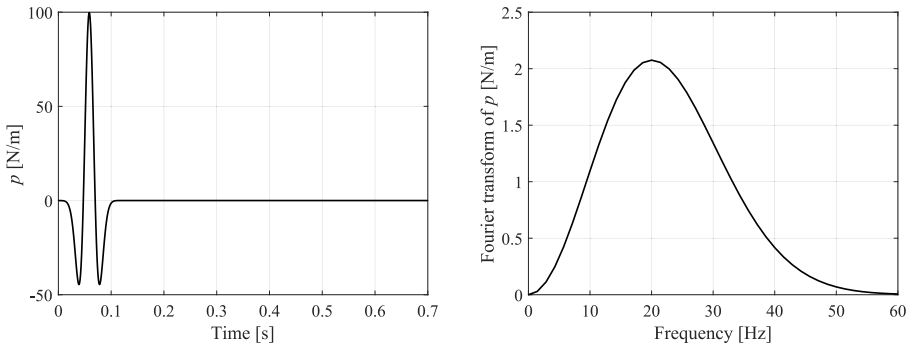
To generate training data for machine learning simulations, we solve the forward problem with elliptical or triangular voids of varying numbers and shapes in each iteration. We consider a layered elastic medium surrounded by a 5 m thick PML on the left and right sides in the x-direction and 10 m thick at the bottom side of the y-direction. The interior domain is 40 m long and 40 m wide in a plane-strain setting discretized by an element size of 1 m. The entire domain is discretized into a structured mesh of 50 × 50 square elements. A schematic of the PML-truncated semi-infinite layered domain with the elliptical and triangular voids can be shown in Fig. 8.

For the first layer, the shear wave velocity is V_s = 150 m s^{-1}, and the dilatational wave velocity is V_p = 300 m s^{-1}. For the second layer, V_s = 200 m s^{-1} and V_p = 400 m s^{-1}. For the third layer, V_s = 250 m s^{-1} and V_p = 500 m s^{-1}. For the fourth layer, V_s = 300 m s^{-1} and V_p = 600 m s^{-1}. The Poisson ratio $\nu = \frac{1}{3}$ and the mass density ρ = 1500 kg/m^3 are considered constant for all the layers. A Ricker-pulse signal with a central frequency of 20 Hz is applied as a point force located at (10, 50) m. The total simulation time for each iteration is 0.7 s with a time step of 0.001 s. We record the measured displacement responses, both u_x and u_y, using 41 receivers evenly placed on the top surface of our domain at a spacing of 1 m.

3.2. Randomizer

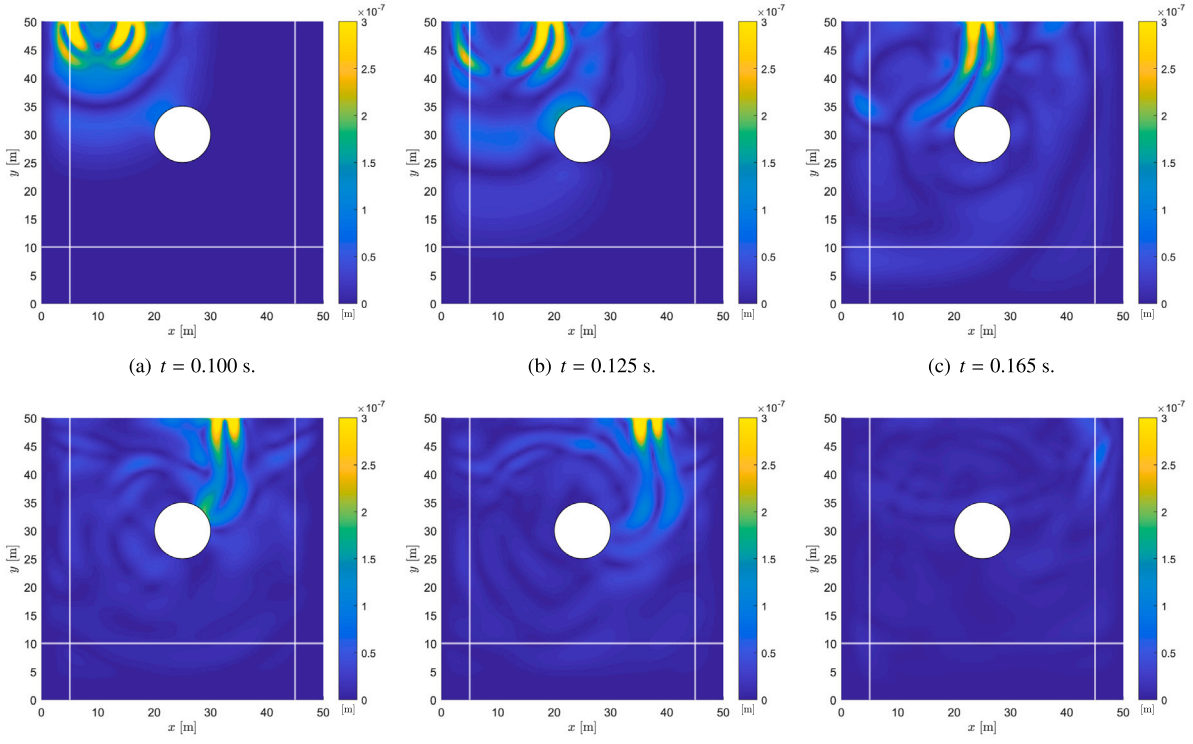
We generate up to two elliptical or triangular voids within the interior domain, each with random sizes and positions. As shown in Fig. 9(a), each elliptical void is centered at (x_0, y_0), with a width of $2m$ and a height of $2n$, oriented at an angle θ. In our datasets, ellipses have widths $2m$ ranging from 8 to 16 m, heights $2n$ ranging from 4 to 10 m, and angles θ ranging from 5° to 175°. As illustrated in Fig. 9(b), each triangle is defined by an initial vertex at (x_0, y_0), an angle α from the x-axis, followed by an internal angle β, and side lengths a and b. In our datasets, triangular voids have angles α ranging from 5° to 85°, β ranging from 20° to 85°, and side lengths a and b ranging from 8 to 16 m.

For bias-free training data, we randomly insert voids over the interior domain. We generate one or two voids in each dataset to have



(a) Time history. (b) Frequency spectrum.

Fig. 5. Time history and frequency spectrum of Ricker-pulse signal.



(a) t = 0.100 s. (b) t = 0.125 s. (c) t = 0.165 s.

(d) t = 0.220 s. (e) t = 0.250 s. (f) t = 0.370 s.

Fig. 6. Time-dependent displacement contour plots obtained from our level-set wave solver at several time instants.

unrelated parameter values. Despite any overlap between voids, we ensure that void elements are not double-counted.

3.3. Input and output layer features

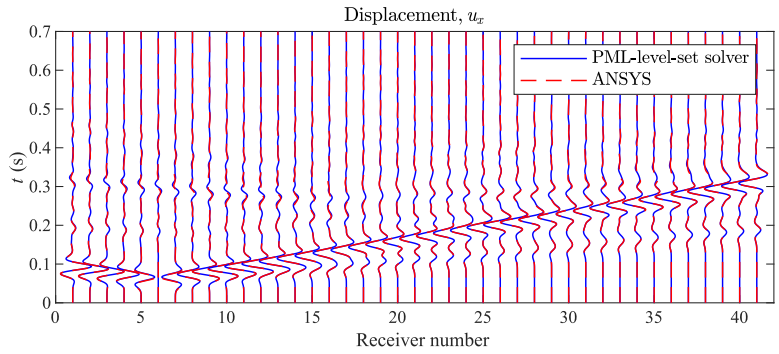
Our data generation process starts with updating the voids' locations, sizes, and numbers for each training sample. After defining the void geometry, we apply an element-wise classification labeling procedure to the computational domain based on the level-set approximation (Pranto et al., 2023; Kim et al., 2024). Every element is then labeled using the level-set approximation, assigning "1" to void elements and "0" to non-void elements. These binary output-feature labels are learned for DCNN to classify each element individually.

The use of element-wise labeling is an important part of our framework. Instead of representing voids using predefined geometric parameters (e.g., size, shape, angle), our labeling process assigns a discrete class to every element independently. This removes any assumptions about the void's shape or number and enables the DCNN to learn a fully flexible, shape-agnostic mapping from wave measurements to the spatial distribution of voids. In effect, the labeling converts the imaging task into a dense classification problem, allowing the network to reconstruct voids of arbitrary geometry, including complex, irregular, or previously unseen shapes.

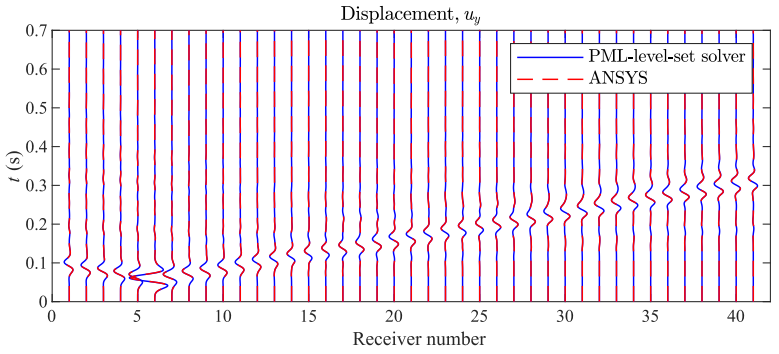
At the same time, our in-house wave solver captures the displacement field of the propagating wave responses at receiver locations for each randomly generated domain. These displacement-time histories constitute the input-layer features. Paired with the element-wise labels, they form a comprehensive dataset that enables the DCNN to learn the nonlinear relationships between wave patterns and the underlying void distribution.

3.4. Data preparation

A total of 18,000 training datasets are generated, comprising three subsets of 6000 datasets corresponding to cases with 0, 1, and 2 voids, respectively. The datasets with one void are further divided into 3000



(a) u_x.



(b) u_y.

Fig. 7. Displacement-time histories recorded at top-surface receivers are compared between our PML–level-set solver and the ANSYS solution.

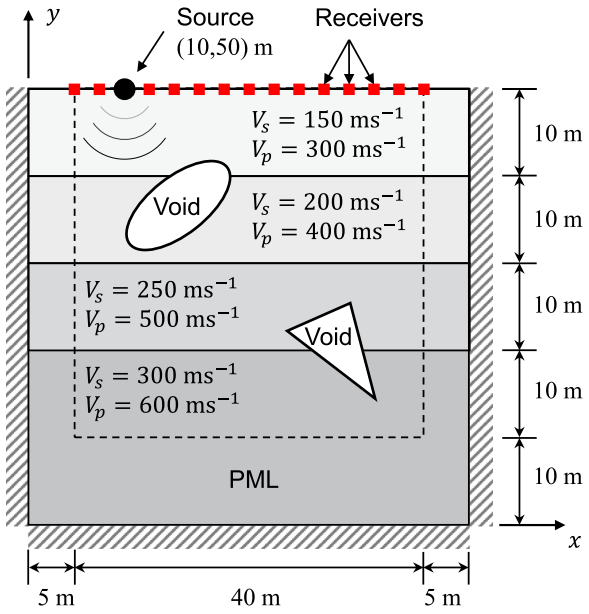


Fig. 8. A PML-truncated semi-infinite layered domain with the elliptical and triangular voids.

datasets with elliptical voids and 3000 datasets with triangular voids. The datasets with two voids are split into 2000 with two elliptical voids, 2000 with two triangular voids, and 2000 with one elliptical void and one triangular void. In our study, we divide the dataset into three subsets: training, validation, and test sets. We ensure that each subset has an equal representation of void categories, specific voids with different combinations of elliptical and triangular shapes, to prevent bias toward any particular shape. The training set is used to train the neural network, while the validation set is used to evaluate its generalization ability on unseen data.

To facilitate effective training and interpretation of results, we standardized the displacement responses, which varied significantly (typically on the order of 10^{-n}, with n varying widely across the dataset). This standardization, achieved using a simple normalization technique, scaled the displacement responses to fall within the range of $[-0.5, 0.5]$. The normalization is described by the following equation:

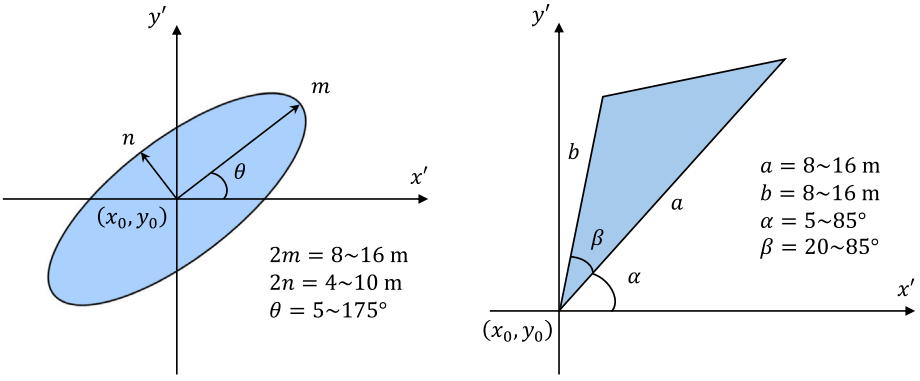
$$A^n_{ijk} = \frac{A_{ijk} - A^{\text{mean}}_{\text{train}}}{A^{\max}_{\text{train}} - A^{\min}_{\text{train}}}, \tag{4}$$

where A^n_{ijk} is the normalized displacement response, A_{ijk} is the original displacement response, and $A^{\text{mean}}_{\text{train}}$, $A^{\max}_{\text{train}}$, and $A^{\min}_{\text{train}}$ represent the mean, maximum, and minimum displacement responses of the training set, respectively. This normalization process facilitates more stable training of the neural network and improves the interpretability of the results. Additionally, this standardized range ensures consistency when applying the trained model to the test dataset, thereby enhancing the model's generalization performance and its applicability to unseen data (Maharjan et al., 2022).

4. Training and evaluation

4.1. Deep Convolutional Neural Network (DCNN)

In this study, we employ a DCNN, shown in Fig. 10, to identify voids within the PML-truncated domain by utilizing wave measurement data. The network architecture consists of multiple one-dimensional (1D) convolutional layers, each followed by a batch normalization layer to improve training stability and convergence (Maharjan et al., 2024). Each convolutional layer–batch normalization pair is regarded as a single convolutional block.



(a) Elliptical void. (b) Triangular void.

Fig. 9. The parameters of the elliptical and triangular voids for random data generation.

Several hyperparameters that govern the model's representational capacity and training efficiency are optimized within predefined search spaces (Table 1). The activation function is selected between Rectified Linear Unit (ReLU) (Nair and Hinton, 2010) and LeakyReLU (Maas et al., 2013) to control the nonlinearity and mitigate vanishing-gradient issues. The kernel initializer is chosen among Glorot normal, Glorot uniform, He normal, and He uniform to regulate the initial weight distribution and stabilize the optimization process (Glorot and Bengio, 2010; He et al., 2015). The number of convolutional layers and the number of filters per layer determine the model depth and feature extraction capability, while the kernel size defines the receptive field in the temporal domain. The learning rate is adjusted to balance convergence speed and stability, while optimizers such as Adam (Adam et al., 2014), Nadam (Dozat, 2016), and RMSprop (Graves, 2013) control the gradient update strategy. An extensive hyperparameter search is performed to identify configurations that minimize the binary cross-entropy loss on the validation dataset, ensuring optimal generalization performance.

The best configuration obtained in this study through the hyperparameter tuning process, together with the corresponding search ranges, is summarized in Table 1. In the final architecture, the output of the last convolutional block is passed through a flatten layer, which converts the learned feature maps into a 1D vector. This vector is then passed to a dense layer with a sigmoid activation function, yielding an element-wise probability value that indicates whether each of the 2500 background elements in the computational domain corresponds to a void or a non-void element.

To visualize the predicted void distribution, the resulting 1D probability vector is reshaped into a 2D array (50 × 50 elements) whose resolution matches the background discretization of the computational domain. For instance, the probability value in the first component of the vector indicates that of the pixel, located at the first row (from top to bottom) and first column (from left to right), in the 2D map. Likewise, Component 2 of the vector maps to the 2D map's pixel at the row 1 and the column 2, and Component 51 of the vector corresponds to the 2D map's pixel at the row 2 and the column 1 (see Fig. 10(e)). This 1D-to-2D redistribution is a standard practice in many numerical fields. For example, in finite element analysis, the solution is typically stored as a 1D vector but is mapped back to 2D or 3D fields for contour visualization. Similarly, in our framework, reshaping the element-wise probability vector into a 2D grid enables the irregular void profile to be displayed as a spatial damage map.

The probabilistic predictions produced by the DCNN are further refined through a post-processing procedure to obtain coherent clusters of predicted void elements. Each 50 × 50 probabilistic map is smoothed using a Gaussian filter to compute local mean and variance, and then converted to a Z-score to identify strong probability values. Elements with Z-scores higher than a prescribed threshold ($Z > 0.8$) and probabilistic amplitudes exceeding 50% of the local maximum are initially selected as defect regions for potential clusters. Finally, the detected clusters are binarized by classifying each element as void or non-void, resulting in the predicted clusters of void elements.

We use the binary cross-entropy loss function, shown in Eq. (5), for computing the loss of the binary probabilistic output, and utilize an evaluation metric (see Eq. (7)) to monitor convergence. Additionally, we implement callbacks to stop the neural network training if performance does not improve over five consecutive epochs.

$$\mathcal{L} = -\frac{1}{M}\sum_{j=1}^{M}\left(\frac{1}{N}\sum_{i=1}^{N} Y_i \log(P_i) + (1 - Y_i)\log(1 - P_i)\right)_j, \tag{5}$$

In this equation, Y_i denotes the target label for the ith element, where "1" and "0" represent void and non-void elements, respectively. N represents the number of elements, M the size of the training set, and P_i the model's estimated void probability for element i, ranging from 0 to 1.

4.2. Evaluation metrics

This section presents the variation of evaluation metrics throughout the training process. The following conventional evaluation metrics can be used to assess the performance of the CNN predictions for a binary classification problem:

$$\text{Accuracy} = \frac{\text{TP} + \text{TN}}{\text{TP} + \text{FN} + \text{TN} + \text{FP}} \times 100\ (\%), \tag{6a}$$

$$\text{Precision} = \frac{\text{TP}}{\text{TP} + \text{FP}} \times 100\ (\%), \tag{6b}$$

$$\text{Recall} = \frac{\text{TP}}{\text{TP} + \text{FN}} \times 100\ (\%), \tag{6c}$$

$$\text{F1-score} = \frac{2 \times \text{Precision} \times \text{Recall}}{\text{Precision} + \text{Recall}} \times 100\ (\%), \tag{6d}$$

where TP, TN, FP, and FN denote the counts of true-positive, true-negative, false-positive, and false-negative classifications within a specific dataset. In detail, TP is counted when the CNN accurately identifies void elements as void, TN when non-void elements are correctly identified as non-void, FP when non-void elements are incorrectly classified as void, and FN when void elements are mistakenly classified as non-void.

Although metrics—Precision, Recall, and F1-score—are available for assessing the performance of a typical binary classification problem, only the Accuracy can consider both TP and TN. Namely, the Accuracy can evaluate the extent to which the method identifies both TP and TN classes, while the others cannot. However, if the dataset exhibits a severe imbalance between positive and negative samples,

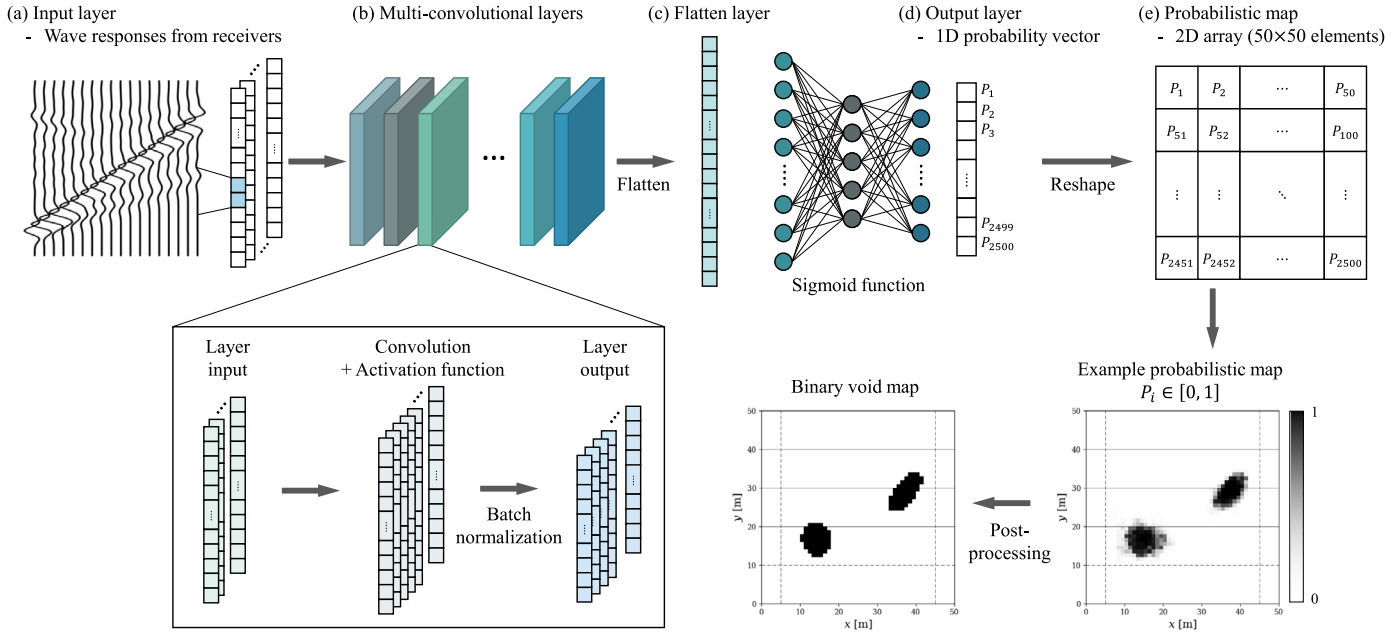


Fig. 10. DCNN architecture.

standard Accuracy may be misleading as it is dominated by the majority class (i.e., negative class). To address the limitation of the standard Accuracy, we utilize an alternative evaluation metric—Class-Weighted Accuracy (CWA)—that better reflects performance across both classes. Specifically, we define

$$\mathrm{CWA} = \frac{w_p \cdot \mathrm{TP} + w_n \cdot \mathrm{TN}}{w_p \cdot (\mathrm{TP} + \mathrm{FN}) + w_n \cdot (\mathrm{TN} + \mathrm{FP})} \times 100\%, \tag{7}$$

where w_p and w_n denote the class-based weights. These weights are defined as the inverse of the class proportions,

$$w_p = \frac{\mathrm{TP} + \mathrm{FP} + \mathrm{TN} + \mathrm{FN}}{\mathrm{TP} + \mathrm{FN}}, \quad w_n = \frac{\mathrm{TP} + \mathrm{FP} + \mathrm{TN} + \mathrm{FN}}{\mathrm{TN} + \mathrm{FP}}. \tag{8}$$

This weighted formulation reduces the dominance of the majority class and ensures that both positive and negative classes contribute comparably to the overall performance metric.

We evaluate the CNN's performance on two test sets: (*i*) in-distribution test datasets generated by our level-set wave solver, with results presented in Section 5, and (*ii*) out-of-distribution test datasets from ANSYS software, which include complex-shaped voids unseen during training. Using these datasets, we also performed several additional evaluations, including a set of parametric analyses and a comparative study with FWI. The corresponding results are detailed in Section 6.

5. Numerical results on in-distribution test data

5.1. Training history of the DCNN model

We demonstrate the convergence of the binary cross-entropy loss function over 27 epochs. As illustrated in Fig. 11(a), the loss value initially starts high for both training and validation datasets and gradually converges to a low minimum. Similarly, Fig. 11(b) shows that CWA begins at a low value and increases steadily before stabilizing over time.

Two key observations are worth noting. First, the training and validation curves closely align, indicating that the model effectively learns feature relationships without overfitting the training data. Second, as the training stabilizes—meaning the model stops learning new patterns due to repeated exposure to the training data—the training process halts as intended by the callback function. This mechanism prevents overfitting and enhances generalization performance.

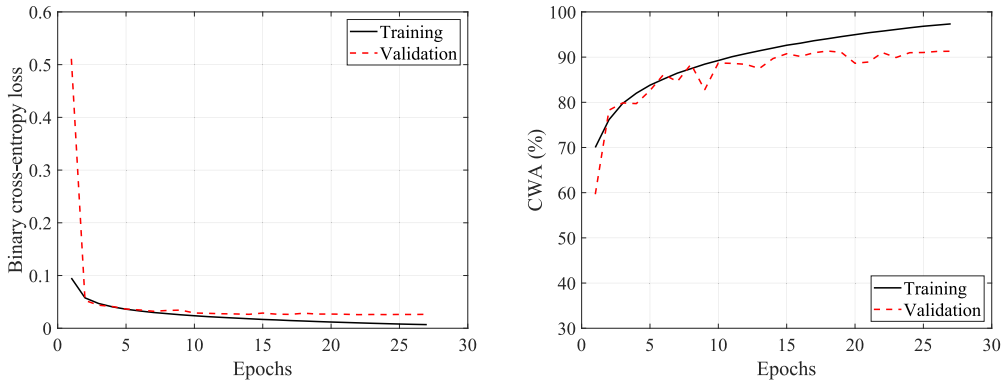
This model is trained with the best hyperparameter configuration obtained by the automated hyperparameter tuning method. The search space for hyperparameters and their best values are illustrated in Table 1.

5.2. Performance on test datasets

For the initial numerical evaluation, we test our CNN on datasets generated by the level-set forward wave solver. The CNN predicts three types of void domains: (*i*) no voids, (*ii*) a single void, and (*iii*) two voids. The results are summarized in Table 2. Based on Table 2, the overall class distribution is highly imbalanced, with void elements representing only a small portion of the computational domain, as discussed in Section 4.2.[1] Because of this imbalance, the standard Accuracy remains high (over 95%), even when some void elements are not correctly identified. Conversely, a number of incorrectly classified void elements (i.e., FN) can reduce Recall and consequently lower the F1-score to a greater extent than Accuracy. Although these metrics are not affected by class imbalance in the same way as Accuracy, they do not account for TN and therefore do not provide a fully comprehensive evaluation. Because of these characteristics, we adopt CWA as the primary evaluation metric, which offers a more balanced representation of performance across void and non-void elements.

Firstly, for domains without voids, the CNN achieves perfect accuracy, as shown in Fig. 12. This success is attributed to the balanced dataset design and training strategy, which included an equal number of samples for each domain type: no voids, one void, and two voids (6000 datasets per type). This distribution ensures that the CNN can effectively learn and distinguish between the different domain types. As a result, the displacement responses and corresponding void element mapping are consistent across all samples, leading to well-learned feature relationships between the responses and the void domains.

[1] Aside from such class imbalance discussed herein, data imbalance is also common in practice for subsurface imaging because field surveys typically provide only a limited number of defect samples (Tao et al., 2025) and can be mitigated through techniques such as data augmentation and semi-supervised learning to address the significant imbalance in the dataset (Zeng et al., 2025).



(a) Binary cross-entropy loss function. (b) CWA.

Fig. 11. Convergence curves of the loss function and CWA over 27 epochs.

Table 1
Optimal hyperparameter configuration of the DCNN trained on 41 receivers.

Hyperparameter	Search space	Best configuration
Activation function	ReLU, LeakyReLU	ReLU
Kernel initializer	Glorot normal, Glorot uniform, He normal, He uniform	Glorot normal
Number of convolutional layers	1–10 (step: 1)	9
Number of filters per layer	10–200 (step: 10)	[160, 90, 120, 90, 170, 190, 70, 190, 20]
Kernel size	10–100 (step: 10)	70
Learning rate	10^{-5}, 10^{-4}, 10^{-3}, 10^{-2}	10^{-5}
Batch size	16, 32, 64	32
Optimizer	Adam, Nadam, RMSprop	RMSprop

Next, we evaluate the CNN's performance in predicting domains with a single void. In the best prediction scenario (Fig. 13), the CNN accurately reconstructs the elliptical void, capturing its shape with high confidence, yielding a CWA of 100%. Even in the worst prediction scenario (Fig. 14), where the CNN reconstructed a nearly triangular shape with slightly smoothed or imperfectly defined vertices, it successfully captured the general location, shape, and presence of a single void.

We then examine our CNN's predictions for domains with two voids. Generally, the CNN model struggles with predicting domains that have two voids. We believe this difficulty can be attributed to several factors, including the location of the wave source, the placement of the receivers, and the existence of wave-absorbing boundary conditions used in this problem. In the best prediction (Fig. 15) of the problem domain with two voids, the CNN produces a reasonable reconstruction of two voids, capturing their locations, sizes, and shapes, yielding a CWA of 99.04%. However, in the worst prediction scenario (Fig. 16), the CNN model achieves a CWA value of 75.94%. In the considered problem setting, when both elliptical and triangular voids are vertically aligned, the amplitude of the wave portion recorded at the top-surface receivers — specifically, the component reflected from a lower void in a vertically aligned pair and subsequently attenuated or reverberated by the upper void — is much smaller than that reflected from a single void. Consequently, the proposed DCNN has limited ability to recognize the relationship between the top-surface wave responses and the presence of a lower void in vertically aligned configurations. This explains why the reconstruction accuracy for the lower void is noticeably lower than that for an isolated void. It is also worth noting that, due to the use of the PML, there are no reflections from the side or bottom boundaries of the computational domain. Without these reflections–which could otherwise illuminate the voids–the waves reflected from the lower void reach the surface receivers with insufficient amplitude to effectively inform the neural network.

6. Numerical results on out-of-distribution test data

We conduct further numerical tests to evaluate the CNN's performance using out-of-distribution test datasets generated by ANSYS Mechanical. By leveraging measured wave responses from this third-party software, which models void boundaries more accurately than our forward wave solver, we effectively reduce the risk of an inverse crime in our analysis. Namely, this method simulates a scenario where the inversion simulator processes experimental data governed by real-world physics. The out-of-distribution test includes cases with elliptical or triangular voids, as well as complex-shaped voids such as hourglass-shaped or cave passage-shaped voids, which were not included in the training dataset.

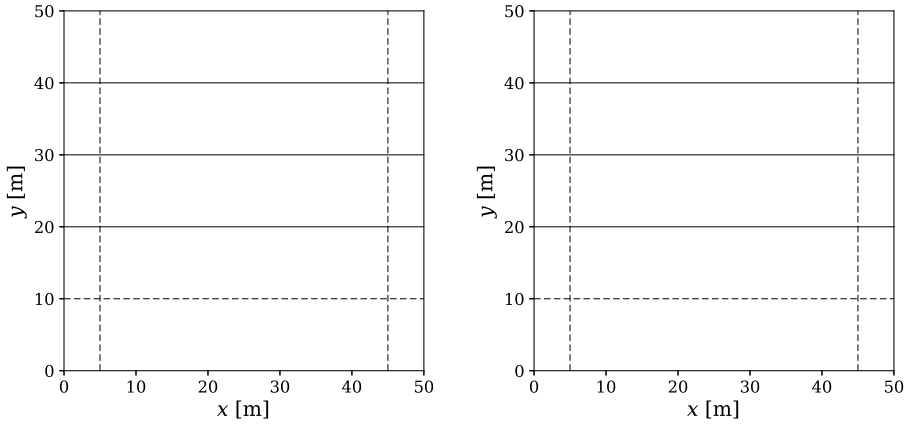
Using these ANSYS-generated out-of-distribution datasets, we conduct a series of sensitivity analyses to determine how different sources of modeling variability shape the CNN's predictions. The following subsections address receiver-related uncertainties and spacing, material variability, changes in excitation frequency and void size/depth, and cases involving a larger number of targeted voids than those used for training. A comparative study with full-waveform inversion (FWI) is also provided.

6.1. Performance with respect to the uncertainty and configurations of receivers

Receiver-related factors represent one of the major contributors to variability in practical measurements. In this section, we consider three such factors: (*i*) measurement noise, (*ii*) random tilting of receivers, and (*iii*) receiver spacing. The model is evaluated with respect to each of the aforementioned factors, and the corresponding results are presented in the following subsections.

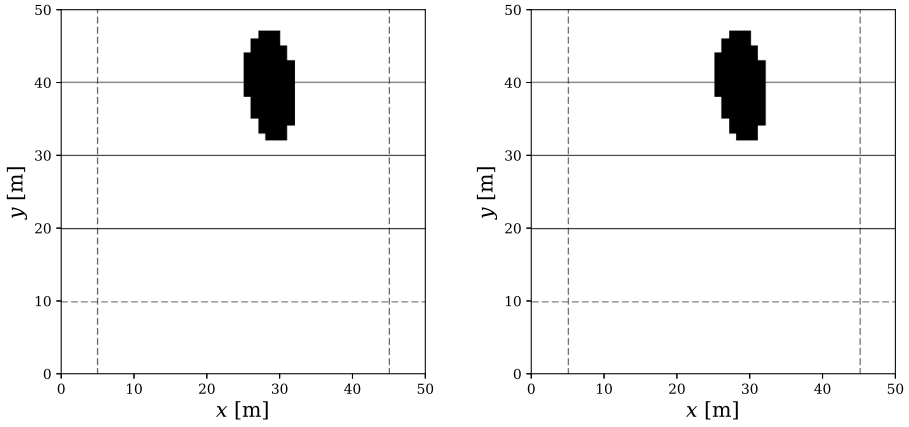
6.1.1. Effect of noise

In this section, we evaluate the CNN's performance on out-of-distribution test data with Gaussian noise levels of 0%, 2%, and 4%. Two training scenarios are considered: (1) a model trained using input data augmented with random Gaussian noise with amplitudes uniformly sampled between 1% and 5%, and (2) a model trained using



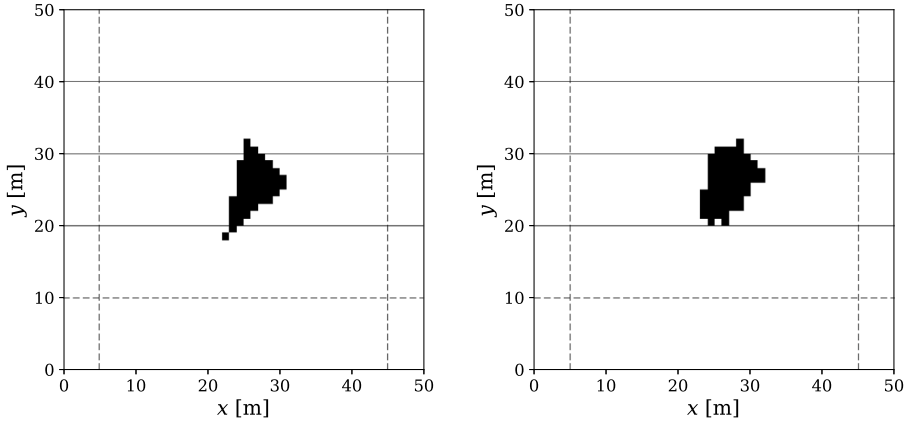
(a) Target domain. (b) Predicted domain.

Fig. 12. Prediction for zero target void in the test dataset.



(a) Target domain. (b) Predicted domain.

Fig. 13. Best prediction for one target void in the test dataset.

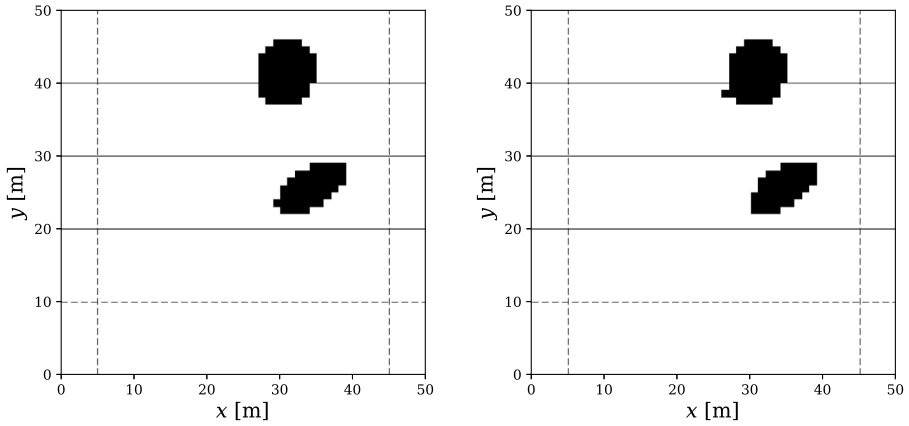


(a) Target domain. (b) Predicted domain.

Fig. 14. Worst prediction for one target void in the test dataset.

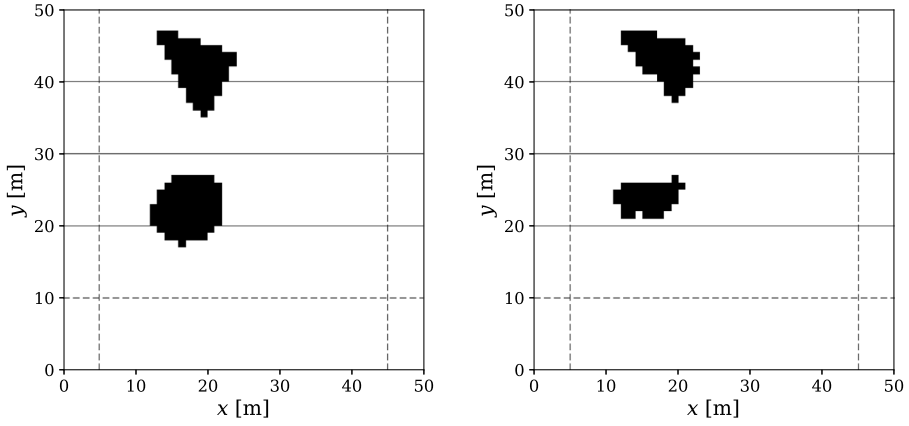
Table 2
Evaluation metrics for the CNN on test datasets.

Predictions	TP	TN	FP	FN	Accuracy (%)	Precision (%)	Recall (%)	F1-score (%)	CWA (%)
Fig. 12 - Prediction (no voids)	0	2500	0	0	100				100
Fig. 13 - Best prediction (one void)	83	2417	0	0	100	100	100	100	100
Fig. 14 - Worst prediction (one void)	35	2442	4	19	99.08	89.74	64.81	75.27	82.33
Fig. 15 - Best prediction (two voids)	105	2392	1	2	99.88	99.06	99.06	98.13	99.04
Fig. 16 - Worst prediction (two voids)	77	2341	12	70	96.72	86.52	52.38	65.25	75.94



(a) Target domain. (b) Predicted domain.

Fig. 15. Best prediction for two target voids in the test dataset.



(a) Target domain. (b) Predicted domain.

Fig. 16. Worst prediction for two target voids in the test dataset.

noise-free input data. The hyperparameters for the best configuration for the retraining with the noise are omitted for brevity.

Figs. 17, 18, and 19 show that CNNs trained with random Gaussian noise consistently outperform those trained without noise when evaluated on out-of-distribution test data containing Gaussian noise. When the CNN is trained without noise, its CWA value drops sharply as the noise level increases in all Examples 1–7, with the largest decrease (32%) occurring in Example 6.

The training process with added noise helps the network learn convolutional filters that not only capture the key signal features associated with the targeted cavity but also smooth out noisy measurement signals. Thus, this numerical experiment indicates that training with noisy measurement data is essential for handling realistic, noise-contaminated experimental conditions.

6.1.2. *Random tilting of receivers' orientation*

In this experiment, we have tested the trained CNN presented in Section 6.1 (noise free, tilting free, 41 receivers) on out-of-distribution test data while measurement data are affected by the tilting of receivers' orientation. To this end, three tilt conditions are considered: 0°, ±5°,

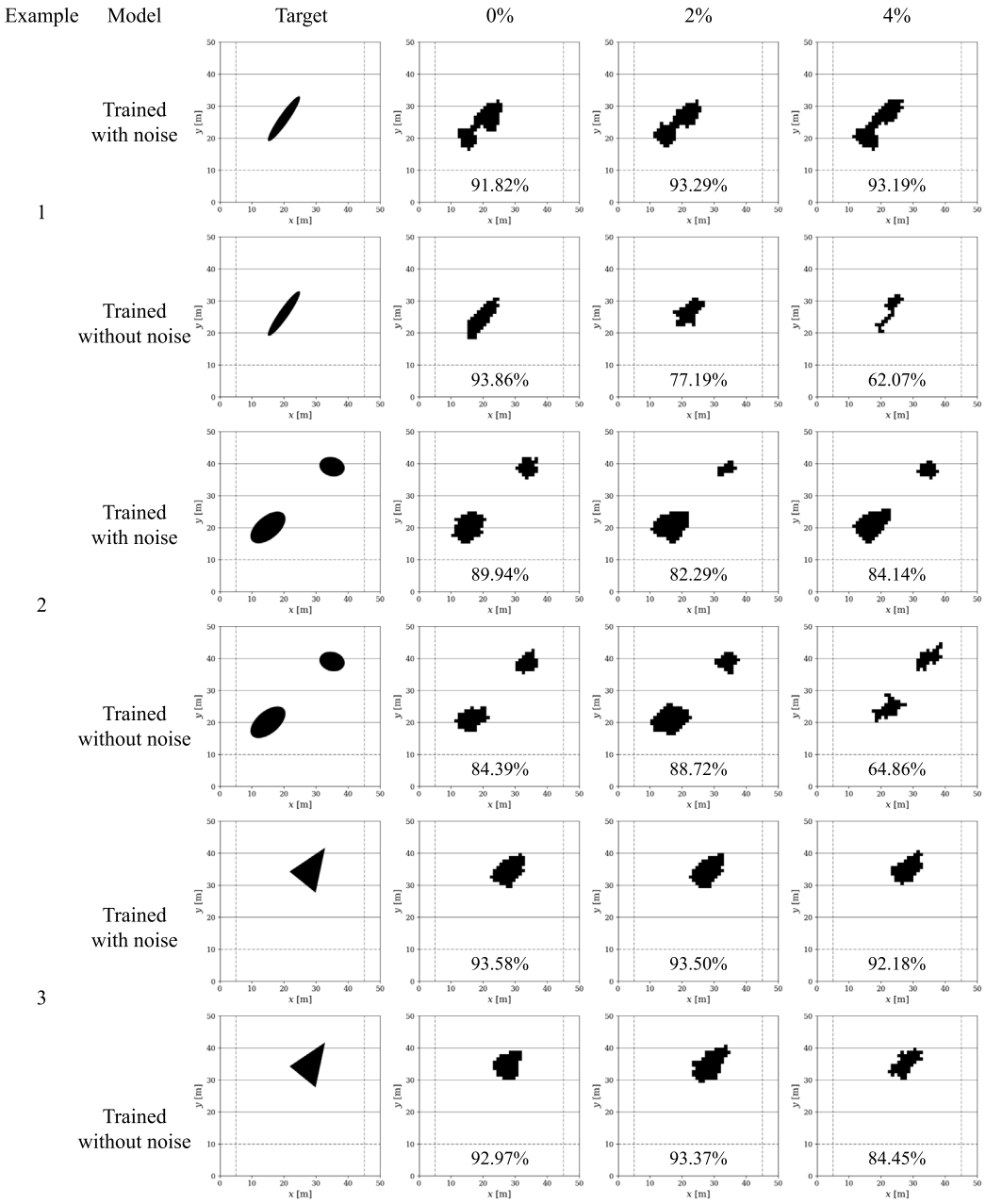


Fig. 17. Targets and predictions for Examples 1 to 3 under different Gaussian noise levels of 0%, 2%, and 4%, along with the corresponding CWA values. Results are compared between noise-free and noise-trained models.

±10°. For the tilted cases, each receiver is randomly assigned, via uniform distribution, a tilt angle between −5° and +5°, or between −10° and +10°, in both the x- and y-directions. For each tilted receiver, the measured displacement components are rotated according to the assigned tilt angle ψ:

$$u_x^{\text{tilted}} = u_x \cos\psi - u_y \sin\psi, \tag{9a}$$

$$u_y^{\text{tilted}} = u_x \sin\psi + u_y \cos\psi, \tag{9b}$$

which represents the coordinate transformation induced by the receiver's orientation. This configuration allows the model to evaluate how variations in receivers' orientation affect the overall imaging performance. Visual results shown in Fig. 20 indicate that, in general, the random tilt angle between −10° and +10° reduces the CWA only by 1 to 2%. This finding suggests that, in practical experimental settings, random tilting of receivers on the ground surface may not significantly degrade imaging accuracy.

6.1.3. Receiver spacing

We investigate the effect of the number of receivers on the model performance. The receiver spacings are 1, 2, 4, 8, and 16 m, corresponding to 41, 21, 11, 6, and 3 receivers, respectively. For receiver spacings of 2, 4, 8, and 16 m, the DCNN is newly retrained and an independent hyperparameter search is performed for each configuration. For receiver spacing of 1 m, we use the baseline model whose hyperparameters are presented earlier in Table 1. The hyperparameters

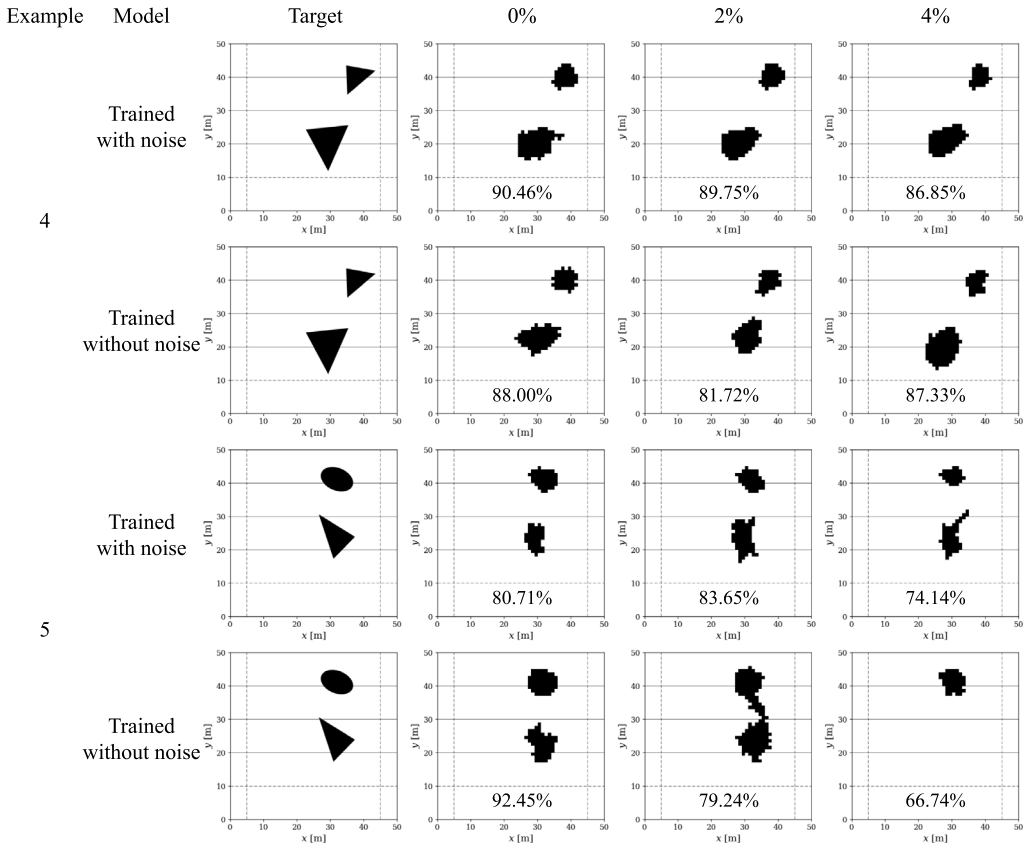


Fig. 18. Targets and predictions for Examples 4 and 5 under different Gaussian noise levels of 0%, 2%, and 4%, along with the corresponding CWA values. Results are compared between noise-free and noise-trained models.

corresponding to the best configuration used for retraining are omitted here for brevity.

Fig. 21 presents the target domains and the corresponding predictions for all seven out-of-distribution test examples obtained under different receiver spacings, along with their associated CWA values. In general, smaller receiver spacing results in higher CWA values, indicating improved reconstruction performance. To quantify this trend more clearly, Fig. 22 summarizes the CWA performance across all spacing configurations and additionally depicts the mean CWA curve.

In our configuration, the dominant wavelength at the receiver layer (i.e., the top layer) is approximately $\lambda = 7.5$ m. As shown in Figs. 21 and 22, receiver spacings smaller than the dominant wavelength still yield high-quality reconstructions. The performance begins to decline as the spacing approaches λ and becomes noticeably degraded once the spacing exceeds this wavelength. According to prior studies in wave-based imaging, spatial sampling finer than approximately half of the dominant wavelength is generally required to resolve scattering features (Shi et al., 2021). In our case, a spacing of 4 m (approximately $\lambda/2$) or smaller yields consistently strong reconstruction performance, with average CWA values above 85%. In contrast, a spacing greater than $\lambda/2$ (i.e., 8 m and 16 m) results in reduced reconstruction performance, with the average CWA dropping to 80% or lower. However, such reduction is less substantial in Examples 1 and 3, with single, simple-shaped targets, than the others. Overall, these observations suggest that using receiver spacing smaller than approximately $\lambda/2$ at the top layer is desirable for maintaining reconstruction quality.

6.2. Material uncertainty in background strata

In this section, we evaluate how variations of material properties in background strata affect the CNN's performance on out-of-distribution test datasets. To generate test datasets, the shear wave velocity V_s and the dilatational wave velocity V_p in each of the four layers are independently perturbed by random factors uniformly sampled from the range of −5% to +5%. Test samples are generated based on the two representative void types from Examples 1 and 2. The void geometry is kept unchanged while wave velocities in background strata are randomized. A total of 100 test samples is generated for each example while all other modeling parameters remain fixed.

Fig. 23 shows the distributions of CWA values for the two examples. For Example 1, the baseline CWA without material variation is 93.86%. When the material properties are varied, a few lower CWA values appear, but higher values in the 95%–100% range are also observed. Overall, the distribution remains concentrated near high CWA values. For Example 2, the baseline CWA is 84.39%. Under material variations, the distribution shifts slightly downward and becomes wider, with additional samples appearing at lower CWA ranges than Example 1.

6.3. Generalization to samples with more voids than those for training

To examine how the network performs on configurations beyond those used in training, we also evaluate it on samples containing three voids. Fig. 24 shows two representative cases. Although the training data include only one- and two-void domains, the network produces reasonably accurate predictions for both samples of three targeted

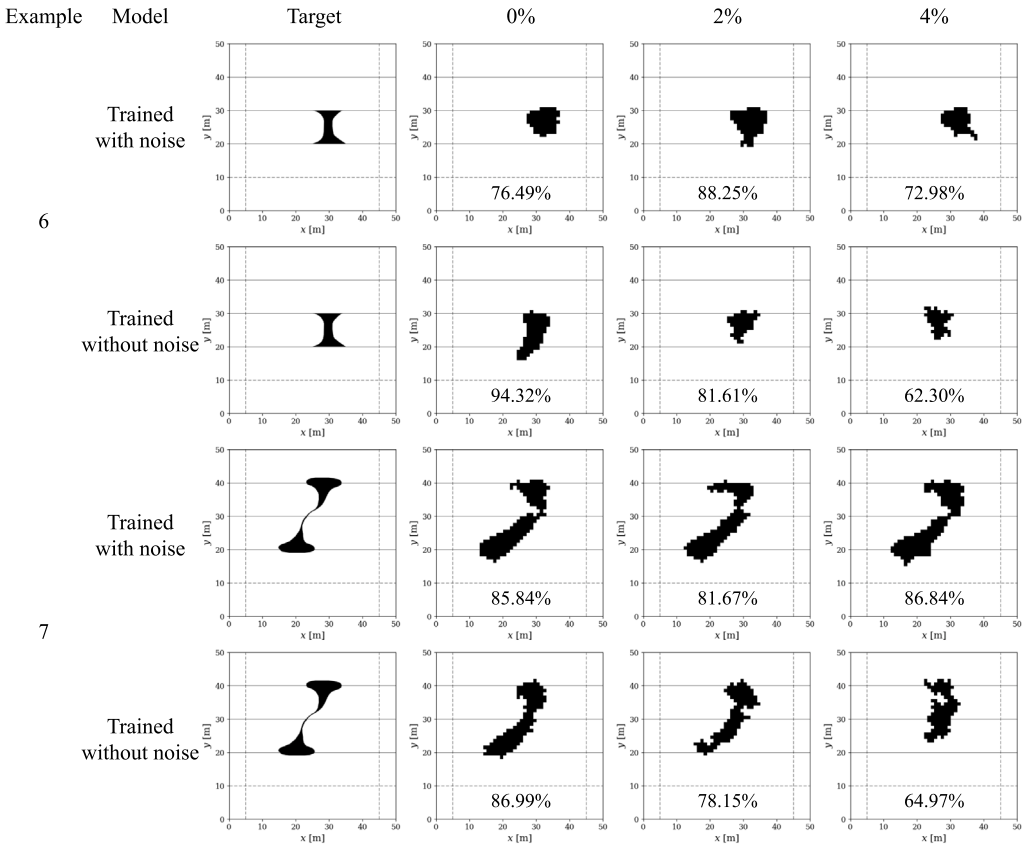


Fig. 19. Targets and predictions for Examples 6 and 7 under different Gaussian noise levels of 0%, 2%, and 4%, along with the corresponding CWA values. Results are compared between noise-free and noise-trained models.

voids. The approximate locations and overall arrangements of the three voids are consistently recovered, and the predicted voids agree well with the targets in terms of spatial distribution. Although the CWA values are around 70%, which is relatively low, the overall correspondence between the targets and predictions remains satisfactory across the tested examples.

6.4. Study on the minimum detectable void size and maximum detectable depth

In our numerical setting, the minimum detectable void size is fundamentally constrained by the wavelength of the input wavefield. The baseline model uses a 20 Hz Ricker pulse for shear-wave velocities of 200–300 m/s in the second to fourth layers, the associated wavelength ranges from approximately 10 m to 15 m. In wave-based imaging, features larger than roughly half the wavelength generally produce resolvable scattering patterns (Shi et al., 2021). Accordingly, voids smaller than about 5 to 7.5 m in the second to fourth layers are not expected to generate sufficiently informative signatures in our forward simulations.

To further examine detectability, we evaluated circular voids with diameters of 4–8 m placed at depths of 10, 20, and 30 m (Fig. 25). Voids of 4–5 m were not detected at any depth under the baseline frequency. Voids of 7–8 m were consistently detected across all tested depths, whereas 6 m voids were detectable at 10 m and 20 m but not at 30 m. This reduction in detectability at larger depths is sensible because deeper layers are associated with longer effective wavelengths and much weaker scattered-wave amplitudes at the surface receivers, making smaller voids more difficult to resolve. Overall, the numerical results meet the aforementioned expectation. In realistic 3D geomechanical media, wave attenuation — arising from both material damping and geometrical spreading — would further diminish the scattered-wave amplitude, making deep and small voids even more challenging to detect.

Regarding the maximum detectable depth, the 20 Hz baseline wavefield allows our model to detect 7–8 m voids down to 30 m in this idealized 2D undamped elastic setting. However, this depth should not be interpreted as a practical limit. In realistic 3D damped solids, frequency-dependent attenuation governs the penetration depth, and achieving reliable detection at deeper depths typically requires enhanced illumination strategies. For example, phased-array source configurations can improve the energy delivery to deeper regions, potentially enabling void detection beyond the depths considered here. Establishing a generalizable detectable-depth limit under such conditions would require a comprehensive analysis across frequencies, attenuation levels, and source configurations, and we consider this an important direction for future work.

6.5. Comparative study with FWI

In this section, we compare the DCNN-based predictions with wave velocity fields reconstructed using a PML-endowed FWI method (Kucukcoban et al., 2019; Kim et al., 2025). To illustrate the imaging characteristics, Fig. 26 shows the results for two representative examples (Examples 1 and 2). Both methods—our DCNN and the FWI—use the same wavefield measurements as input. Moreover, the initial guess of the material properties for the FWI is identical to those of the layered profile used in the forward simulations that generated the

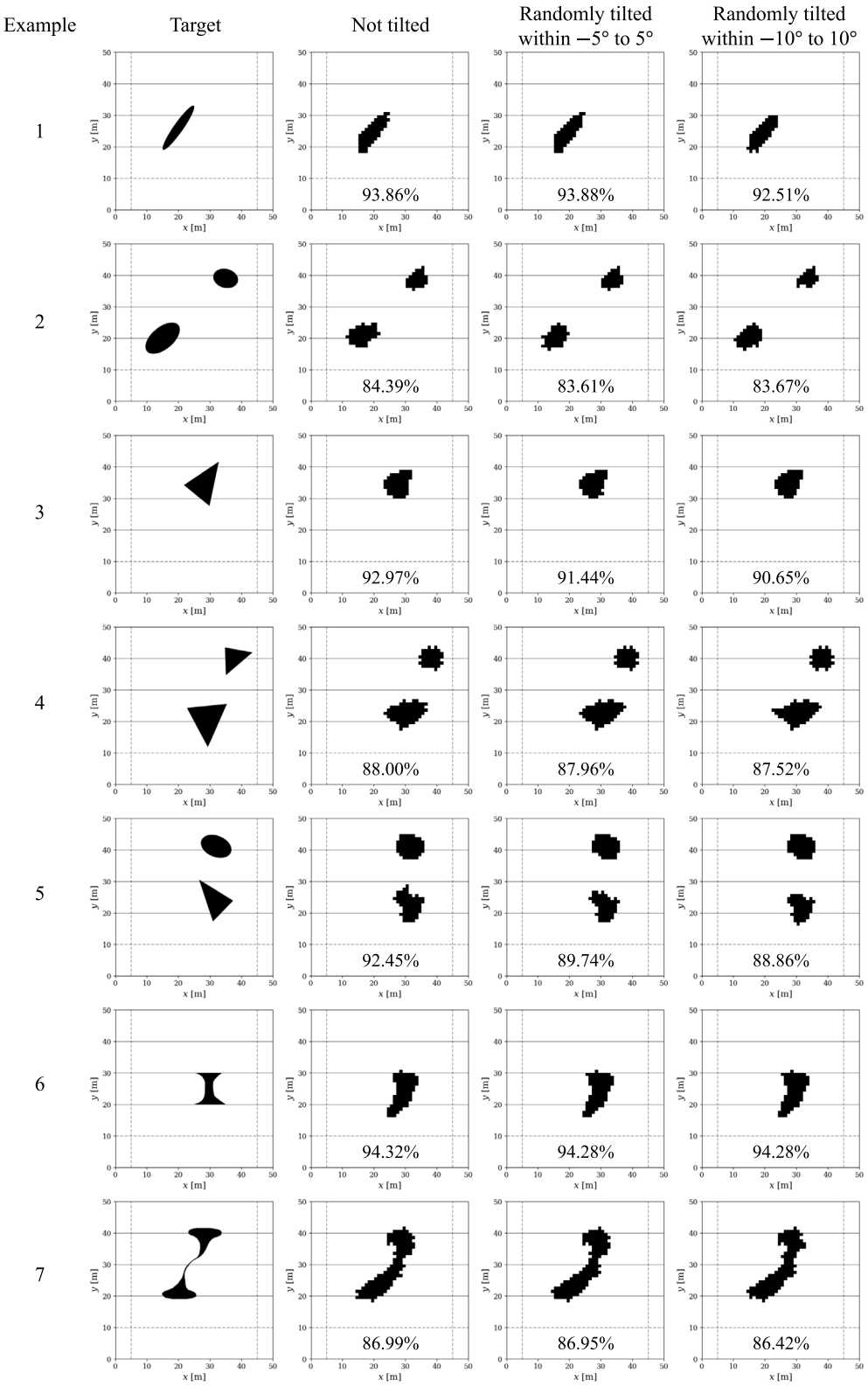


Fig. 20. Targets and predictions obtained under different receiver tilting conditions, along with the corresponding CWA values.

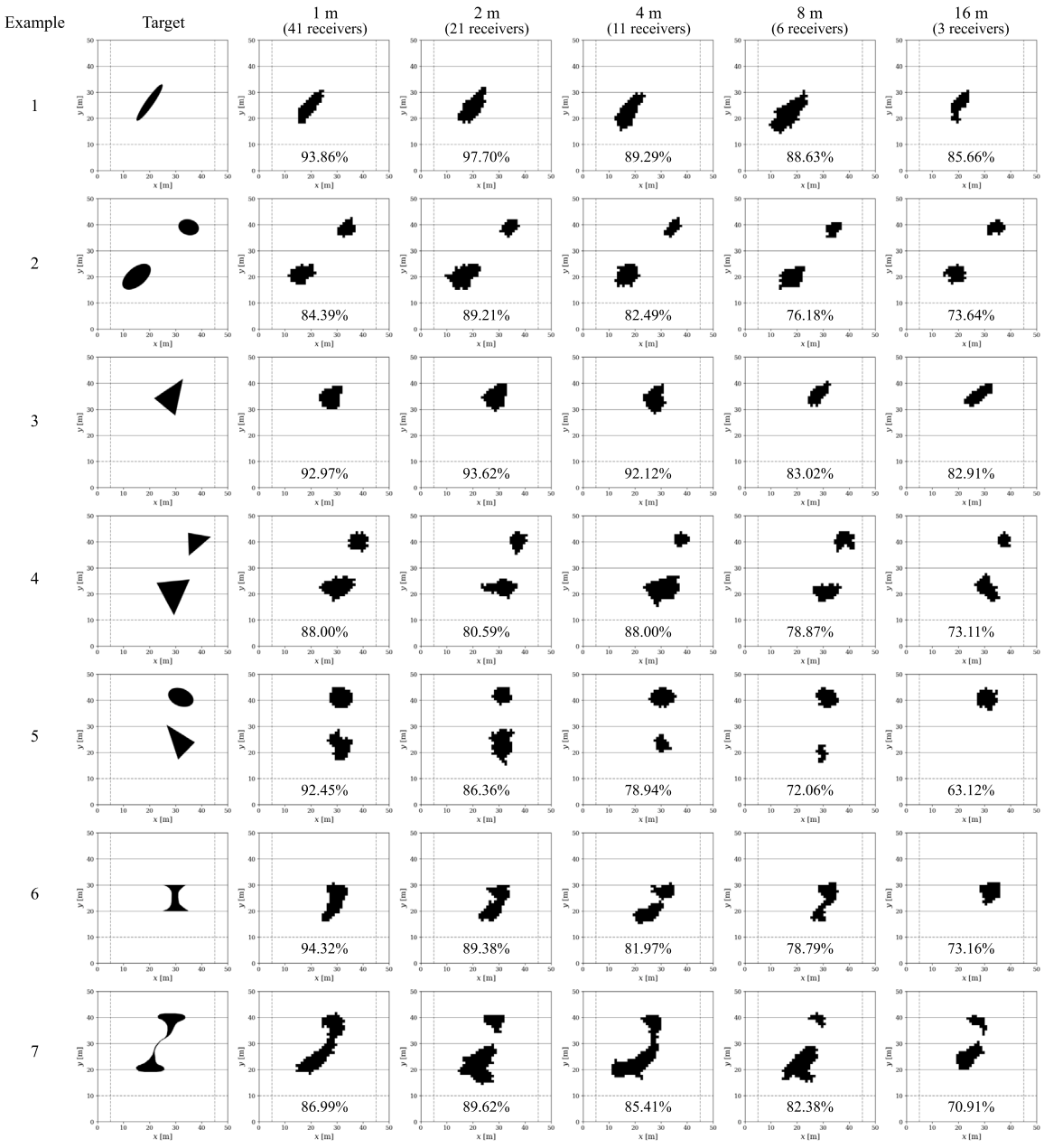


Fig. 21. Targets and predictions obtained under different receiver spacings, along with the corresponding CWA values.

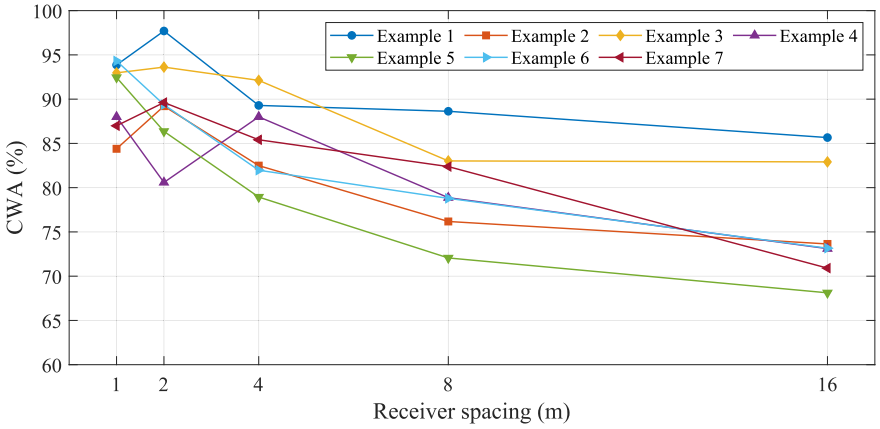
DCNN training data. During the FWI process, the mass density and the Poisson ratio are kept identical to those for the DCNN model. Thus, this numerical example does not bias either method and enables a fair comparison.

In Example 1, the DCNN-based prediction clearly identifies the void at the correct location, while the FWI reconstruction shows a reduced-velocity region in the similar area but with a slightly shallower depth representation than the target. Example 2 shows that the DCNN-based prediction successfully identifies both voids at the correct locations, whereas the FWI reconstruction reveals a reduced-velocity region only around the upper void and fails to capture the lower one. This may be because updating the velocity profile around the upper void (especially, its upper-left tip) is more effective than doing so around the lower void in terms of reducing the misfit functional during the FWI process.

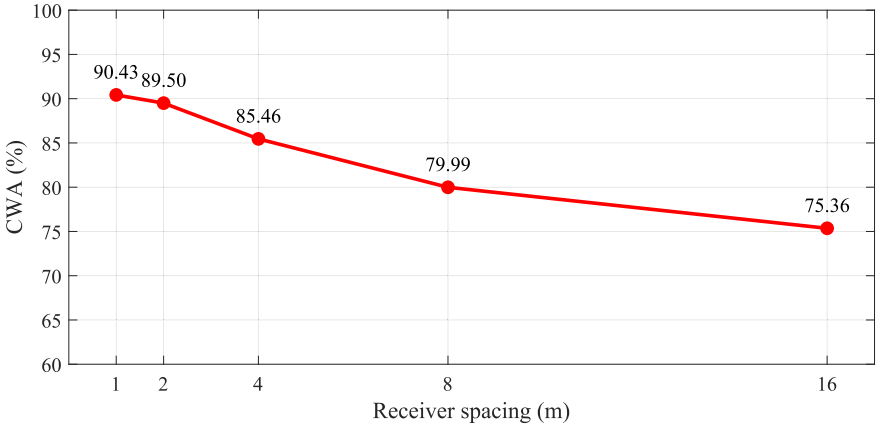
The FWI reconstruction represents the reduced-velocity regions in a smooth and continuous manner, which can make it less straightforward to distinguish individual voids, especially in cases involving deeper or multiple voids. It confirms our literature review on the limitation of the FWI as it can only reconstruct a targeted void as areas of low wave speeds and hardly delineates the void. In contrast, the DCNN provides discrete classifications for each element, allowing the void locations to be shown more explicitly and clearly than FWI in these cases.

7. Conclusion

This study shows a DCNN method to detect and image the locations, shapes, and sizes of voids within a PML-truncated layered domain using elastodynamic waves. The main contributions of this study are summarized as follows:

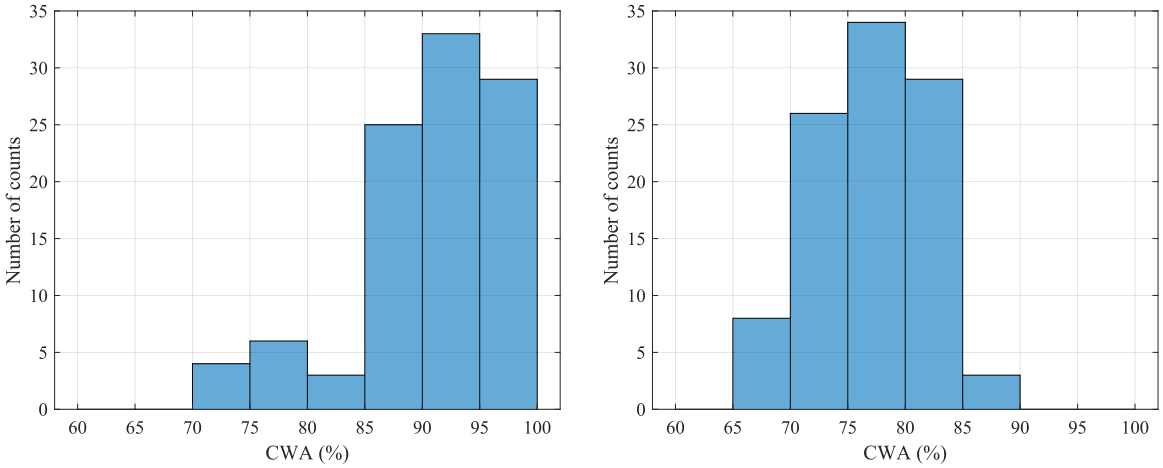


(a) CWA values for Examples 1 to 7.



(b) The mean value of CWA across receiver spacings.

Fig. 22. CWA performance under different receiver spacings, presented for Examples 1 to 7 and the overall mean.



(a) Example 1. (b) Example 2.

Fig. 23. Distributions of CWA values under material uncertainty in Examples 1 and 2.

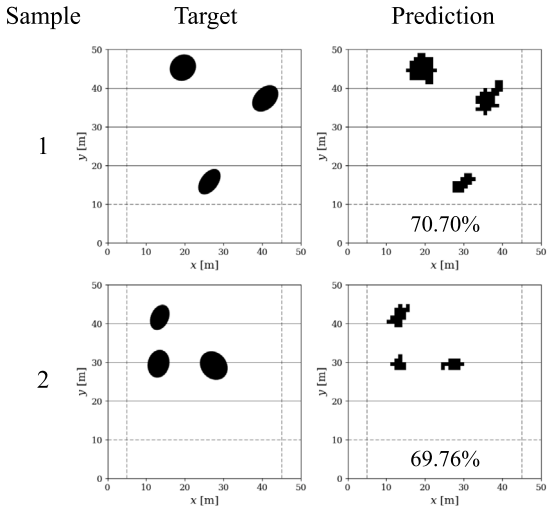


Fig. 24. Targets and predictions for representative samples containing three voids, along with the corresponding CWA values.

- A level-set elastodynamic forward solver is employed to simulate wavefields within domains containing voids, without re-meshing the computational domain. A semi-infinite layered model using CFS-PML is then used to generate high-fidelity training wavefield data while minimizing reflections from the truncated boundaries.
- An element-wise classification formulation is adopted, in which each background element is independently predicted as void or non-void. This formulation reconstructs voids through clusters of high-probability elements and avoids requiring prior assumptions about void characteristics, allowing the CNN to identify voids of arbitrary shapes, numbers, locations, and sizes from elastodynamic surface responses.

The proposed CNN is evaluated on in-distribution test datasets generated by a PML-level-set wave solver to predict three types of void domains: no voids, one void, and two voids. For domains without voids, the model achieved perfect accuracy due to the balanced training with equal samples for each domain type. In predicting single voids, the CNN demonstrated strong performance, accurately reconstructing the void's general location and shape. For domains with two voids, lower performance is observed when the voids are vertically aligned. The limitations are attributed to wave source positioning, receiver locations, and the absorbing boundary conditions, which hindered wave illumination into a deeper void beneath a shallower one. Nevertheless, the model successfully identified the approximate location and number of voids, demonstrating its practical potential.

We also conduct out-of-distribution tests using a third-party simulator, in which the model explicitly defines void boundaries, avoiding an inverse crime and achieving greater accuracy than our forward wave solver. The CNN demonstrates strong generalization capabilities, reasonably predicting the location, shape, and size of various voids, even when these voids are smaller or more complex than those in the training data. The main findings from the out-of-distribution tests are summarized as follows:

- The model remains stable under receiver-related uncertainties such as measurement noise and random tilting of receivers. Noise-trained model, in particular, substantially improves stability under noisy conditions.
- Even when the receiver layout becomes sparse, the CNN continues to provide reasonable predictions.
- The model exhibits stable performance under the examined uncertainties in material properties of the background strata.
- The CNN generalizes to cases outside the training conditions (i.e., considering only up to two voids) and, thus, is able to reconstruct samples containing three voids.
- Whereas FWI reconstructs the voids through smooth reduced-velocity regions, the CNN provides a more spatially localized representation than the FWI that makes the void locations appear more clearly.

The numerical experiments also demonstrate the limitation of the presented method. First, the predicted void boundaries tend to be rounded, which results in, for instance, the sharp corners of the triangular targets being inaccurately reconstructed as curved edges. Second, the DCNN struggles to associate surface responses with lower voids in vertically aligned configurations, leading to reduced reconstruction accuracy. Lastly, the presented method requires a-prior information of background stratigraphy that can be obtained from SASW, MASW, or FWI material inversion analysis. Once the strata's material profile is determined, the presented ML process can clearly image hidden voids within the strata.

Nomenclature

Symbol	Comment
x, y, z	Horizontal, vertical, and anti-plane directions, respectively
$V(x, y)$	Enrichment function
$\mathbf{u}$	Displacement field
u_x, u_y	x- and y-direction displacements
$\phi(x, y)$	Local shape function
N_n	The number of nodes within an element
$F(t)$	Ricker-pulse wave source signal over time t
V_s	Shear wave velocity
V_p	P-wave velocity
f	Central frequency of the Ricker pulse
$2m$, $2n$	Width and height of elliptical void
θ	Angle of elliptical void
a, b	Side lengths of triangular void
α	Angle from the x-axis of triangular void
β	Internal angle of triangular void
A^n_{ijk}	Normalized displacement response
A_{ijk}	Original displacement response
$A^{\text{mean}}_{\text{train}}$, $A^{\text{max}}_{\text{train}}$, and $A^{\text{min}}_{\text{train}}$	Mean, maximum, and minimum displacement responses, respectively
$\mathcal{L}$	Loss function
Y_i	Target label for the ith element
$P(Y_i)$	Predicted probability of the ith element being void
N	Number of elements
M	Number of training datasets
TP, TN	Counts of true-positive and true-negative classifications
FP, FN	Counts of false-positive and false-negative classifications
ψ	Tilting angle of the receiver
u_x^{tilted}, u_y^{tilted}	x- and y-direction displacements in tilted coordinate direction
λ	Wavelength

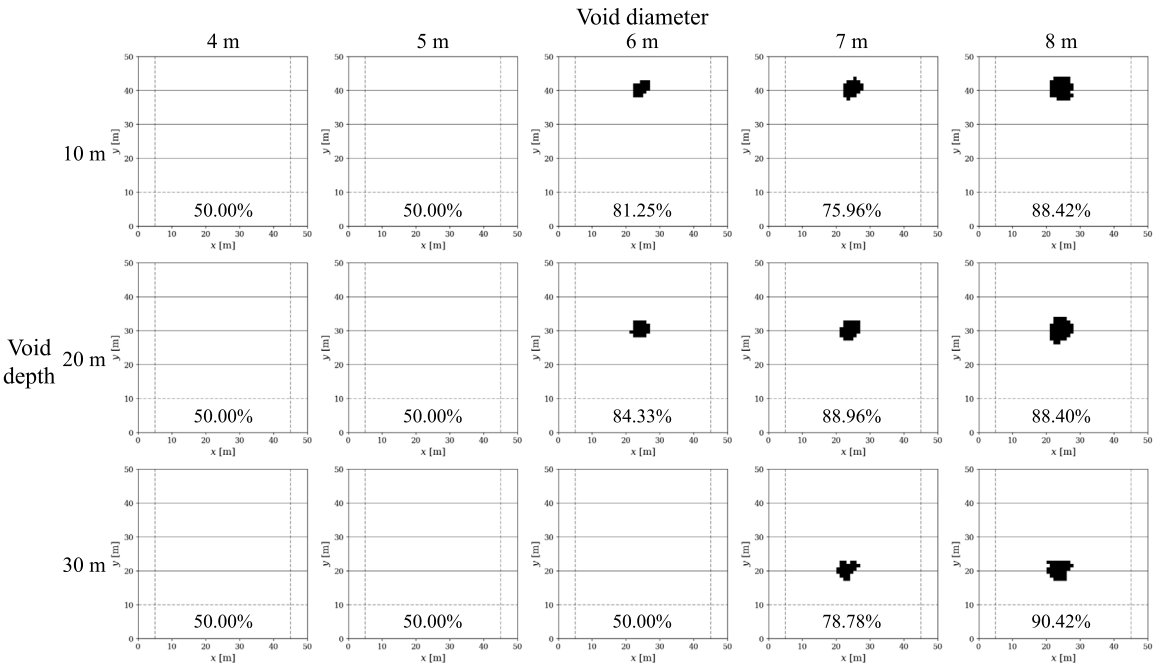


Fig. 25. Detection results for voids of varying diameters placed at different depths, along with the corresponding CWA values.

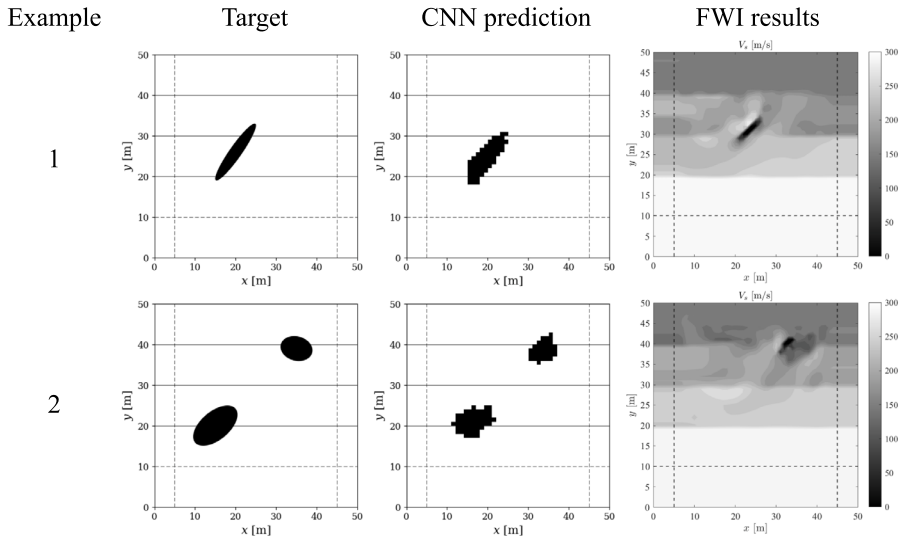


Fig. 26. Comparison of the target voids with the DCNN-based predictions and the FWI-reconstructed shear-wave velocity fields using the same wavefield measurements for Examples 1 and 2.

CRediT authorship contribution statement

Boyoung Kim: Writing – review & editing, Writing – original draft, Visualization, Validation, Software, Methodology, Investigation, Formal analysis, Data curation, Conceptualization. **Shashwat Maharjan:** Writing – original draft, Visualization, Validation, Software, Methodology. **Bruno Guidio:** Software, Data curation. **Jacob Thomas:** Software. **Fazle Mahdi Pranto:** Software. **Chanseok Jeong:** Writing – review & editing, Writing – original draft, Validation, Supervision, Software, Resources, Project administration, Methodology, Investigation, Funding acquisition, Formal analysis, Conceptualization.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgments

This material is based upon work supported by the National Science Foundation, United States under Award CMMI-2053694. Any opinions, findings, conclusions, or recommendations expressed in this material

are those of the authors and do not necessarily reflect the views of the National Science Foundation. The authors also acknowledge the technical and financial support of the Automotive Research Center (ARC) from the University of Michigan (UofM) Ann Arbor, MI.

Data availability

Data will be made available on request.

References

Aaker, O.E., Raknes, E.B., Arntsen, B., 2021. Elastodynamic full waveform inversion on GPUs with time-space tiling and wavefield reconstruction. J. Supercomput. 77, 2416–2457.

Adam, K.D.B.J., et al., 2014. A method for stochastic optimization. 1412. arXiv preprint arXiv:1412.6980.

Berenger, J.-P., 1994. A perfectly matched layer for the absorption of electromagnetic waves. J. Comput. Phys. 114 (2), 185–200.

Bharti, A.K., Pal, S., Priyam, P., Kumar, S., Srivastava, S., Yadav, P.K., 2016. Subsurface cavity detection over patherdih colliery, jharia coalfield, India using electrical resistivity tomography. Environ. Earth Sci. 75, 1–17.

Chatzi, E.N., Hiriyur, B., Waisman, H., Smyth, A.W., 2011. Experimental application and enhancement of the XFEM–GA algorithm for the detection of flaws in structures. Comput. Struct. 89 (7), 556–570.

Dai, Q., Lee, Y.H., Sun, H.-H., Ow, G., Yusof, M.L.M., Yucel, A.C., 2022. DMRF-UNet: A two-stage deep learning scheme for GPR data inversion under heterogeneous soil conditions. IEEE Trans. Antennas and Propagation 70 (8), 6313–6328.

Dashwood, B., Gunn, D., Curioni, G., Inauen, C., Swift, R., Chapman, D., Royal, A., Hobbs, P., Reeves, H., Taxil, J., 2020. Surface wave surveys for imaging ground property changes due to a leaking water pipe. J. Appl. Geophys. 174, 103923.

de Assis, F.M., Gomes, G.F., 2021. Crack identification in laminated composites based on modal responses using metaheuristics, artificial neural networks and response surface method: a comparative study. Arch. Appl. Mech. 91 (10), 4389–4408.

De Giorgi, L., Leucci, G., 2014. Detection of hazardous cavities below a road using combined geophysical methods. Surv. Geophys. 35, 1003–1021.

Dozat, T., 2016. Incorporating nesterov momentum into adam.

Fathi, A., Poursartip, B., Kallivokas, L.F., 2015. Time-domain hybrid formulations for wave simulations in three-dimensional PML-truncated heterogeneous media. Internat. J. Numer. Methods Engrg. 101 (3), 165–198.

Francois, S., Goh, H., Kallivokas, L.F., 2021. Non-convolutional second-order complex-frequency-shifted perfectly matched layers for transient elastic wave propagation. Comput. Methods Appl. Mech. Engrg. 377, 113704.

Giammarinaro, B., Tsarsitalidou, C., Hillers, G., de Rosny, J., Seydoux, L., Catheline, S., Campillo, M., Roux, P., 2023. Seismic surface wave focal spot imaging: numerical resolution experiments. Geophys. J. Int. 232 (1), 201–222.

Glorot, X., Bengio, Y., 2010. Understanding the difficulty of training deep feedforward neural networks. In: Proceedings of the Thirteenth International Conference on Artificial Intelligence and Statistics. JMLR Workshop and Conference Proceedings, pp. 249–256.

Graves, A., 2013. Generating sequences with recurrent neural networks. arXiv preprint arXiv:1308.0850.

Guidio, B., Goh, H., Kallivokas, L.F., Jeong, C., 2024. On the reconstruction of the near-surface seismic motion. Soil Dyn. Earthq. Eng. 177, 108414.

Gupta, M., Khan, M.A., Butola, R., Singari, R.M., 2022. Advances in applications of non-destructive testing (NDT): A review. Adv. Mater. Process. Technol. 8 (2), 2286–2307.

He, K., Zhang, X., Ren, S., Sun, J., 2015. Delving deep into rectifiers: Surpassing human-level performance on imagenet classification. In: Proceedings of the IEEE International Conference on Computer Vision. pp. 1026–1034.

Jiang, S., Wan, C., Sun, L., Du, C., 2022. Flaw classification and detection in thin-plate structures based on scaled boundary finite element method and deep learning. Internat. J. Numer. Methods Engrg. 123 (19), 4674–4701.

Jung, J., Jeong, C., Taciroglu, E., 2013. Identification of a scatterer embedded in elastic heterogeneous media using dynamic XFEM. Comput. Methods Appl. Mech. Engrg. 259, 50–63.

Kallivokas, L., Fathi, A., Kucukcoban, S., Stokoe II, K., Bielak, J., Ghattas, O., 2013. Site characterization using full waveform inversion. Soil Dyn. Earthq. Eng. 47, 62–82.

Kang, J.W., Kallivokas, L.F., 2010. Mixed unsplit-field perfectly matched layers for transient simulations of scalar waves in heterogeneous domains. Comput. Geosci. 14 (4), 623–648.

Kim, B., Kang, J.W., 2019. A time-domain formulation of elastic waves in heterogeneous unbounded domains. Multiscale Sci. Eng. 1, 220–235.

Kim, B., Kang, J.W., 2024. Time domain modeling of elastic waves using a stress-based unsplit-field perfectly matched layer with enhanced numerical stability. Appl. Math. Model. 128, 431–449.

Kim, B., Kang, J.W., Choi, Y.-T., 2025. Two-dimensional characterization of railway subgrade using full-waveform inversion of elastic waves. Int. J. Rail Transp. 1–31.

Kim, B., Maharjan, S., Pranto, F.M., Guidio, B., Schaal, C., Jeong, C., 2024. Convolutional neural network and level-set spectral element method for ultrasonic imaging of delamination cavities in an anisotropic composite structure. Ultrasonics 107254.

Kucukcoban, S., Goh, H., Kallivokas, L.F., 2019. On the full-waveform inversion of lame parameters in semi-infinite solids in plane strain. Int. J. Solids Struct. 164, 104–119.

Kucukcoban, S., Kallivokas, L.F., 2011. Mixed perfectly-matched-layers for direct transient analysis in 2D elastic heterogeneous media. Comput. Methods Appl. Mech. Engrg. 200 (1–4), 57–76.

Li, H., Zhang, J., Ye, M., Cheng, Z., Wu, C., Tian, Y., 2025a. Automated detection of multi-scale voids in airport runways with clutter suppression method from 3D GPR data. Constr. Building Mater. 492, 142779. http://dx.doi.org/10.1016/j.conbuildmat.2025.142779, URL: https://www.sciencedirect.com/science/article/pii/S0950061825029307.

Li, H., Zhang, J., Ye, M., Wang, Q., Wei, M., Zhang, Z., Guo, Y., Tian, Y., Zhang, Y., Wang, C., Xu, Z., 2025b. High-accuracy intelligent detection of centimeter-level voids in cement pavement via ground-penetrating radar. Measurement 256, 118354. http://dx.doi.org/10.1016/j.measurement.2025.118354, URL: https://www.sciencedirect.com/science/article/pii/S0263224125017130.

Liu, L., Shi, Z., Li, S., Peng, M., Tao, F., 2023. A system for inspecting karst voids during construction of cast-in-place pile foundations. Eng. Geol. 320, 107124.

Luo, T.X., Lai, W.W., 2020. GPR pattern recognition of shallow subsurface air voids. Tunn. Undergr. Space Technol. 99, 103355.

Maas, A.L., Hannun, A.Y., Ng, A.Y., et al., 2013. Rectifier nonlinearities improve neural network acoustic models. In: Proc. Icml. Vol. 30, Atlanta, GA, p. 3.

Maharjan, S., Guidio, B., Fathi, A., Jeong, C., 2022. Deep and convolutional neural networks for identifying vertically-propagating incoming seismic wave motion into a heterogeneous, damped soil column. Soil Dyn. Earthq. Eng. 162, 107510.

Maharjan, S., Guidio, B., Jeong, C., 2024. Convolutional neural network for identifying effective seismic force at a DRM layer for rapid reconstruction of SH ground motions. Earthq. Eng. Struct. Dyn. 53 (2), 894–923.

Mirzanejad, M., Tran, K.T., McVay, M., Horhota, D., Wasman, S.J., 2021. Deep void detection with 3D full waveform inversion of surface-based and in-depth source seismic wavefields. Eng. Geol. 294, 106407.

Nair, V., Hinton, G.E., 2010. Rectified linear units improve restricted boltzmann machines. In: Proceedings of the 27th International Conference on Machine Learning. ICML-10, pp. 807–814.

Nasseri-Moghaddam, A., Cascante, G., Phillips, C., Hutchinson, D., 2007. Effects of underground cavities on Rayleigh waves—Field and numerical experiments. Soil Dyn. Earthq. Eng. 27 (4), 300–313.

Newmark, N.M., 1959. A method of computation for structural dynamics. J. Eng. Mech. Div. 85 (3), 67–94.

Niu, S., Srivastava, V., 2022. Ultrasound classification of interacting flaws using finite element simulations and convolutional neural network. Eng. Comput. 38 (5), 4653–4662.

Oldenborger, G.A., Routh, P.S., Knoll, M.D., 2005. Sensitivity of electrical resistivity tomography data to electrode position errors. Geophys. J. Int. 163 (1), 1–9.

Park, M.K., Park, S., Yi, M.-J., Kim, C., Son, J.-S., Kim, J.-H., Abraham, A.A., 2014. Application of electrical resistivity tomography (ERT) technique to detect underground cavities in a karst area of South Korea. Environ. Earth Sci. 71, 2797–2806.

Pranto, F.M., Maharjan, S., Jeong, C., 2023. Level-set and learn: Convolutional neural network for classification of elements to identify an arbitrary number of voids in a 2D solid using elastic waves. J. Eng. Mech. 149 (6), 04023035.

Rahimi, M., Wood, C., Kallivokas, L.F., 2024. A comparative study of using geophysical methods for imaging subsurface voids of various sizes and at different depths. Eng. Geol. 341, 107711. http://dx.doi.org/10.1016/j.enggeo.2024.107711, URL: https://www.sciencedirect.com/science/article/pii/S0013795224003119.

Rahnema, H., Ehsaninezhad, L., Dashti, F., Talebi, G., 2022. Detection of subterranean cavities and anomalies using multichannel analysis of surface wave. Geomech. Geoengin. 17 (1), 206–219.

Saad, R., Nawawi, M., Mohamad, E., 2012. Groundwater detection in alluvium using 2-D electrical resistivity tomography (ERT). Electron. J. Geotech. Eng. 17, 369–376.

Shi, S., Lin, L., Luo, Z., Sun, X., Jin, S., 2021. Resolution enhancement of ultrasonic imaging at oblique incidence by using WTFM based on FMC-AR. Measurement 183, 109798.

Sun, H., Waisman, H., Betti, R., 2014. A multiscale flaw detection algorithm based on XFEM. Internat. J. Numer. Methods Engrg. 100 (7), 477–503.

Tao, Y., Zeng, S., Zhou, K., Sun, H., 2025. Cross-domain feature-enhanced YOLOv8 model for underground defect detection. Reliab. Eng. Syst. Saf. 111400.

Thitimakorn, T., Kampananon, N., Jongjaiwanichkit, N., Kupongsak, S., 2016. Subsurface void detection under the road surface using ground penetrating radar (GPR), a case study in the Bangkok metropolitan area, Thailand. Int. J. Geo-Eng. 7, 1–9.

Tran, K.T., McVay, M., Faraone, M., Horhota, D., 2013. Sinkhole detection using 2D full seismic waveform tomography. Geophysics 78 (5), R175–R183.

Virieux, J., Operto, S., 2009. An overview of full-waveform inversion in exploration geophysics. Geophysics 74 (6), WCC1–WCC26.

Wu, Y., Lin, Y., 2018. InversionNet: A real-time and accurate full waveform inversion with CNNs and continuous CRFs. arXiv preprint arXiv:1811.07875.

Xu, L., Wang, K., Su, Y., He, Y., Yang, J., Yuan, S., Su, Z., 2022. Surface/subsurface crack-scattered nonlinear rayleigh waves: a full analytical solution based on elastodynamic reciprocity theorem. Ultrasonics 118, 106578.

Zeng, S., Wang, G., Sun, H., Tao, Y., Pan, X., 2025. A semi-supervised dual-path model for underground defect detection. Eng. Appl. Artif. Intell. 158, 111493.

Zhang, J.-H., Feng, S.-J., Zheng, Q.-T., Zhang, X.-L., 2025a. Fracture image classification study of clay hydraulic fracturing based on non-destructive testing and machine learning methods. Eng. Geol. 108149.

Zhang, G., Zhao, M., Zhang, J., Du, X., 2024. Prismatic-element SBPML coupled with SBFEM for 3D infinite transient wave problems. Comput. Methods Appl. Mech. Engrg. 427, 117014.

Zhang, G., Zhao, M., Zhang, J., Gao, Z., Du, X., 2025b. A non-convolutional complex-frequency-shifted SBPML for wave problems in 3D unbounded waveguide modellings. Appl. Math. Model. 116277.