Multi-Defect Reconstruction in Nondestructive Testing: An interpretable Neural Network Approach

Hairui LIU[1], Zhi Qian[1], Guangming Zhang[2], Peng Li[1], Shirsendu Sikdar[3], D.Z.LIU[4,†] , Zhenghua QIAN[1,†], Iren Kuznetsova[5]

1. State Key Laboratory of Mechanics and Control for Aerospace Structures, College of Aerospace Engineering, Nanjing University of Aeronautics and Astronautics, Nanjing 210016, China;

2. School of Engineering, Liverpool John Moores University, Liverpool, L3 3AF, UK

3. School of Computing and Engineering, University of Huddersfield, Huddersfield, HD1 3DH, UK

4. School of Engineering, Mathematics and Physics, University of East Anglia, Norwich, NR4 7TJ, UK

5. Kotel'nikov Institute of Radio Engineering and Electronics of RAS, Moscow, 125009, Russia

Abstract: Guided wave tomography (GWT) methods for precise multi-defect reconstruction are crucial for structural health monitoring. In this work, an improved physics-informed wave tomography framework (PIWT) is proposed for the quantitative reconstruction of multiple defects in plates. A Trunk-Branch Network is employed to reconstruct the wave travel time and velocity field by synergizing the waveguide governing equations and the real travel time data from sensors. This approach speeds up the network convergence of loss function which includes the travel time data, its first-order derivatives, and the physical principle of wave equations to constrain the space of parameters for accurate defect reconstruction. Based on simulation data, the results demonstrate that PIWT achieve the highly accurate defect with the errors of 4.25% in position and 5.5% in depth. Also, experimental validations are conducted to demonstrate the feasibility of PIWT with a defect position error of less than 1.7% and depth location error under 15%. Furthermore, Uniform Manifold Approximation and Projection (UMAP) is applied to enable a clear visualization of trajectories representing the defect reconstruction convergence, thereby revealing how incremental sensor data enhance the model's capability to approximate the true solution. This interpretation provides useful insights into the latent dynamics to bridge the gap between the black-box nature of deep neural networks and the need for transparent and explainable AI, ultimately reinforcing confidence in the model's applicability for broader engineering applications.

Key words: Guided wave tomography, interpretability, UMAP, defect reconstruction, Lamb wave

1.Introduction

Corrosion of pipelines and storage tanks poses a significant challenge in industries such as petrochemical and nuclear sectors[1]. The detection and quantification of wall thickness loss due to corrosion are gaining increasing attention. Traditional methods for ultrasonic thickness measurement are often cumbersome and expensive, particularly for hard-to-reach areas [2]. Guided wave tomography, however, offers a promising alternative for estimating the remaining thickness of corroded areas without the need for full surface access [3] . By exploiting the dispersive nature of guided waves, this method enables the reconstruction of thickness maps by inversely analyzing ultrasonic signals captured by transducer arrays surrounding the inspection region.

Travel time tomography, also known as travel time layer imaging, is a widely used technique for guided wave defect reconstruction. This method reconstructs the slowness (the reciprocal of the wave velocity) field of a structure by utilizing the arrival time of direct wave packets within guided wave signals. The reconstructed slowness field is then correlated with the structure's thickness using dispersion curves. However, this method assumes that diffraction effects are negligible and is only valid when the defects are much larger than the wavelength of the guided wave. Additionally, the resolution of this method is limited by the first Fresnel zone width, $\sqrt{L\lambda}$, where L is the distance between the source and the receiver, and λ is the wavelength of the illuminating wavefield[6]. Several algorithms have been adapted and tested, including filtered back-projection (FBP), probabilistic (PRA), and algebraic reconstruction techniques (ART)[7]. FBP is originally developed for medical imaging and requires a circular sensor arrangement, which limits its applicability in extended pipelines[10]. In contrast, PRA and ART algorithms demonstrated greater flexibility and effectiveness in imaging curved structures[11]. However, these two algorithms often struggle to provide accurate reconstructions of defect locations and sizes. More recently, Stylianos Livadiotis et al.[12] introduced a two-step method for evaluating internal corrosion in cylindrical structures using spiral-guided ultrasonic waves. This approach effectively estimates the position and size of corrosion defects, but it assumes a single damage scenario, making it unsuitable for the reconstruction of multiple defects.

Neural networks have long been recognized as black-box, data-driven approaches[13]. With the widespread application of Deep Neural Networks (DNNs) in computer vision, natural language processing, medical diagnosis, and other domains, the contradiction between their exceptional predictive performance and their intrinsic “black-box” nature has become increasingly prominent[14]. Although DNNs demonstrate the superhuman capabilities of tackling complex tasks and jobs, their inherent lack of interpretability in both internal decision-making logic and accurately predicted outputs restrains their reliable deployment in critical areas such as healthcare, autonomous driving, and financial risk management[15]. Therefore, Explainable Artificial Intelligence (XAI) has emerged as a focal research area in recent years and an increasing number of researchers have paid attention to improving the interpretability and introducing directional constraints to neural networks by incorporating physical information. With the increasing demand for interpretable predictions **Error! Reference source not found.**, the need for methods that facilitate these objectives has also grown. Raissi et al. **Error! Reference source not found.** developed a deep learning framework for solving and discovering partial differential equations (PDEs). The physics-informed neural network (PINN), known for its versatility, serves as an effective function approximator with the capability of addressing both direct and inverse problems related to PDEs. This has, to some extent, enhanced the explainability of machine learning models.

Natural datasets have intrinsic patterns, which can be summarized using the manifold distribution principle: the distribution of a class of data is represented by a low-dimensional manifold[18]. Based on this principle, Uniform Manifold Approximation and Projection (UMAP) is utilized for dimensionality reduction of high-dimensional data which are challenging to analyze directly[19]. The intrinsic distance features contained in the high-dimensional space are preserved in the low-dimensional space and visualized. Based on the definition provided in [20] to distinguish and explore the difference between interpretability and explainability, interpretability is primarily linked to the intuition behind the outputs of a model, thus the more interpretable a machine learning system is, the easier it becomes to identify cause-and-effect relationships between inputs and outputs of the system.

This study proposes an improved method called the physics-informed wave tomography (PIWT) network with interpretability, designed to achieve the quantitative reconstruction of multiple defects in structures. The methodology primarily leverages the prior physical information to construct the initial model and loss function of the proposed PIWT. This approach serves to constrain the learning direction of the neural network, ensuring stable and fast convergence towards the global optimal solution during the training process. Meanwhile, PIWT addresses the challenge of prematurely falling into local optima during the reconstruction of multiple defects, enabling the precise reconstruction of defect sizes and depth locations. The structure of this paper is described as follows: Section 2 introduces the PIWT principle, covering both forward modeling and inversion techniques. Section 3 details the setup of finite element simulations, data processing, and reconstruction predictions. Section 4 presents the equipment for in-lab tests and signal processing for experimental data. A systematic model interpretability of the proposed methodology is provided in Section 5, followed by a comprehensive discussion. Finally, Section 6 concludes the study.

2. Theory

To derive the plate-thickness maps from the transmission signals of guided waves, the formulation of an inverse problem is essential. It is also noted that the embedded information from the ultrasonic waveforms should be effectively extracted for the reconstruction of the defect's geometric representation. Therefore, a forward model is initially utilized to describe the interaction between the probing waves and the defect. At the other hand, this model acts as a crucial tool, facilitating the inversion process to reveal the defect's geometric details from the captured ultrasonic signals. The details of both the forward model and the inversion scheme are discussed in the following sections.

2.1. the Forward Problem

The applications of guided wave tomography for defect inspection in flat plates have demonstrated that the scattering of guided waves due to plate thinning can be effectively modeled using an acoustic framework[21]. This model utilizes the dispersion characteristics of Lamb waves, as described in [22] to illustrate the dependence of both phase and group velocities of Lamb waves on the product of signal frequency (f) and plate thickness (d). This critical relationship is illustrated in Figure. 1, derived from the analytical solution to the Rayleigh-Lamb frequency equations[23] for both symmetric (S_i) and antisymmetric (A_i) wave modes that can propagate within an aluminum plate.

For a selected mode and frequency, it is hypothesized that the interaction between the probing guided wave and a region of reduced plate thickness resembles the disturbance caused by an acoustic pressure wave traveling through a medium with a non-uniform sound speed distribution. This variation in sound speed across the defect is systematically captured by evaluating the product of frequency and thickness ($f{\cdot}d$) at various points within the defect. The corresponding phase velocity for each point is then obtained from the mode's dispersion curve, providing a detailed understanding of the defect's depth profile based on acoustic velocity variations.

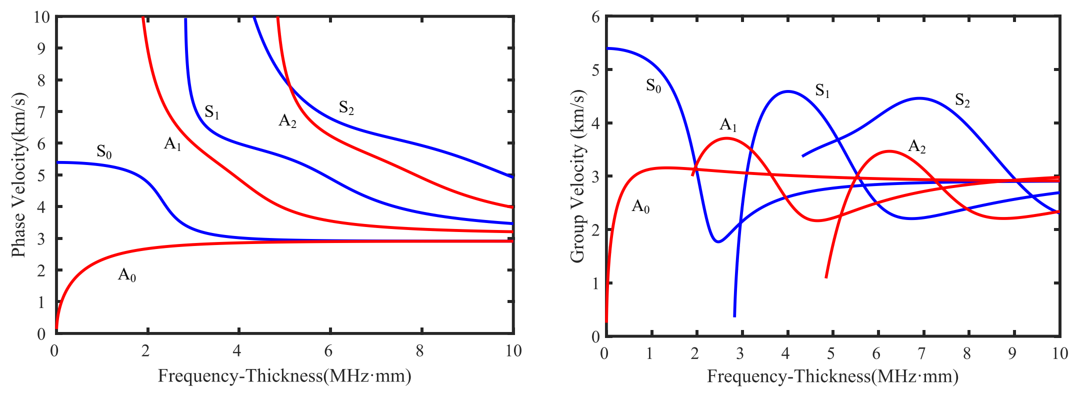


Figure 1. Lamb wave dispersion curves in aluminum plates showing phase velocity (a) and group velocity (b) as functions of the frequency-thickness ($f{\cdot}d$) product for all the propagating modes up to 10 MHz-mm. Blue and red curves refer to the so-called symmetric and antisymmetric modes, respectively.

Further extending this concept to include the plate geometry, it is evidenced that the effect of plate thinning on the propagation dynamics of guided waves can be conceptualized through a phase velocity field, which is defined as follows:

$$c(\mathbf{r}) = c_{ph}[f \cdot d(\mathbf{r})], \quad \forall \mathbf{r} \in \Sigma \tag{1}$$

where $\mathbf{r}$ is the position coordinates $O(x, y, z)$ in three-dimensional (3-D) Euclidean space with Cartesian coordinates. c_{ph} represents the dispersion function of the chosen Lamb mode, and $d(\mathbf{r})$ denotes the distribution of thickness across the plate.

Utilizing the velocity field defined in Eq.1, a two-dimensional acoustic medium is conceptualized. Thus, the scattering of guided waves is simplified into the acoustic scattering that is formulated by a scalar potential $\phi(x, y, f)$ which satisfies the inhomogeneous Helmholtz equation [25]:

$$\nabla^2 \phi(x, y, f) + k_0^2 \phi(x, y, f) = -4\pi O_H(x, y, f)\phi(x, y, f) \tag{2}$$

where $k_0 = 2\pi f / c_0$ is the background wavenumber dependent on the background velocity

$c_0 = c(fd_0)$, with d_0 being a reference thickness, e.g. the thickness of the undamaged plate. $O_H(x, y, f)$ is the object function given by:

$$O_H(x, y, f) = \frac{k_0^2}{4\pi}\left[\left(\frac{c_0(f)}{c(x, y, f)}\right)^2 - 1\right] \tag{3}$$

and serves as the acoustic representation of the loss in plate thickness. $O_H(x, y, f)$ symbolically denotes the effect that essentially vanishes outside the damaged region. For the high-frequency domain, Eq.3 asymptotically aligns with the principle of ray theory within the realm of geometrical acoustics. This alignment is further demonstrated through the Eikonal equation, indicating the essence of wave propagation paths as follows:

$$\left(\frac{\partial T}{\partial x}\right)^2 + \left(\frac{\partial T}{\partial y}\right)^2 = O_e^2(x, y) \tag{4}$$

where the function $T(x, y)$ depicts the travel time of the acoustic wave to the point (x,y) and the object function $O_e(x, y)$ is henceforth defined as the slowness:

$$O_e(x, y) = \frac{1}{c(x, y)} \tag{5}$$

where $c(x, y)$ corresponds to either the phase velocity at the central frequency of the wave pulse or its group velocity at the same frequency. The choice between these velocities depends on the specific aspect of wave propagation under consideration, that is to say, whether the focus is on the propagation of the signal's phase or its energy.

To address the challenges posed by singularities at the point source, the Eikonal equation is transformed by the incorporation of multiplicative coefficients. This modification is achieved by bifurcating the unknown scalar displacement field of waves into two distinct components[26] as follows:

$$T(x, y) = T_0(x, y)\,\tau(x, y) \tag{6}$$

where $T_0(x, y)$ is analytically defined by:

$$T_0(x, y) = \frac{\sqrt{(x - x_S)^2 + (y - y_S)^2}}{c^2(x_S, y_S)} \tag{7}$$

where (x_S, y_S) is the coordinate of the source point.

Substituting Eq.6 into Eq.4, the modified Eikonal equation for solving the forward propagation problem under the consideration of boundary conditions, is obtained as follows:

$$\begin{gathered} T_0^2 \, |\nabla\tau'|^2 + \tau'^2 \left|\nabla T_0\right|^2 + 2T_0\tau'\left(\nabla T_0 . \nabla\tau'\right) = \frac{1}{c^2\left(x, y\right)} \\ \tau'\left(x_s, y_s\right) = 1 \end{gathered} \tag{8}$$

where τ' is the normalized travel time formulated by Eq.9 :

$$\tau'(x, y) = \frac{\tau(x, y) - \tau_{\min}}{\tau_{\max} - \tau_{\min}} \tag{9}$$

where the scaling factors $\tau_{\max}$ and $\tau_{\min}$ are set to 1.3 and 0.9 in this study by many numerical experiments, respectively. To ensure the stability of the network output at the source, τ' is set to 1.

2.2. the Inversion Problem

In this section, the physics-informed wave tomography framework (PIWT) network is proposed for the predictions of defect position and depth location. Its methodology begins with signal acquisition and processing in Fig. 2a, where continuous wavelet transform (CWT) is applied to convert the acquired probe signals into the time-frequency domain. In the time-frequency domain, the time corresponding to the peaks of the signals is extracted as the time of flight (TOF). The TOF values received by all sensors are then assembled into a TOF matrix by M_S X N_S, where M_S represents the number of transmitting probes and N_S denotes the number of receiving probes. The PIWT network is subsequently employed to reconstruct the defect using the TOF matrix and source point location information, yielding the travel time field (*T*) and velocity field (*c*) of the inspected plate. The Eikonal equation is utilized to define the network residuals, thereby constraining the iteration direction of the network in cases of incomplete data. The final physical loss terms are defined as:

$$\begin{aligned} & f := T_0^2 \, |\nabla\tau'|^2 + \tau'^2 \left|\nabla T_0\right|^2 + 2T_0\tau'\left(\nabla T_0 . \nabla\tau'\right) - \frac{1}{c^2\left(x, y\right)} \\ & h := \tau'\left(x_s, y_s\right) - 1 \end{aligned} \tag{10}$$

The PIWT model is based on a Trunk-Branch Network, consisting of two fully connected networks with three hidden layers, each containing 40 neurons, as shown in Fig. 2(b). The Trunk and Branch

Networks are employed separately to predict τ' and c, rather than using a single network with dual outputs. This approach is developed to address the significant differences in the numerical characteristics of τ' and c, preventing the competition between the two predicted values within a single network during the training for network parameters. Moreover, both the zeroth and first-order derivatives of the travel time coefficient are leveraged, effectively incorporating the prior physical knowledge to speed up the network convergence. The Trunk-Branch strategy successfully mitigates the issue of the network becoming trapped in local minima during the reconstruction of multiple defects. The network parameters are optimized using the "Adam" algorithm[27], which estimates the first-order and second-order moments of the gradient, allowing for dynamic adjustment of the learning rate for each parameter. The PIWT neural network are trained on NVIDIA GeForce RTX 3050 GPUs using the TensorFlow[28].

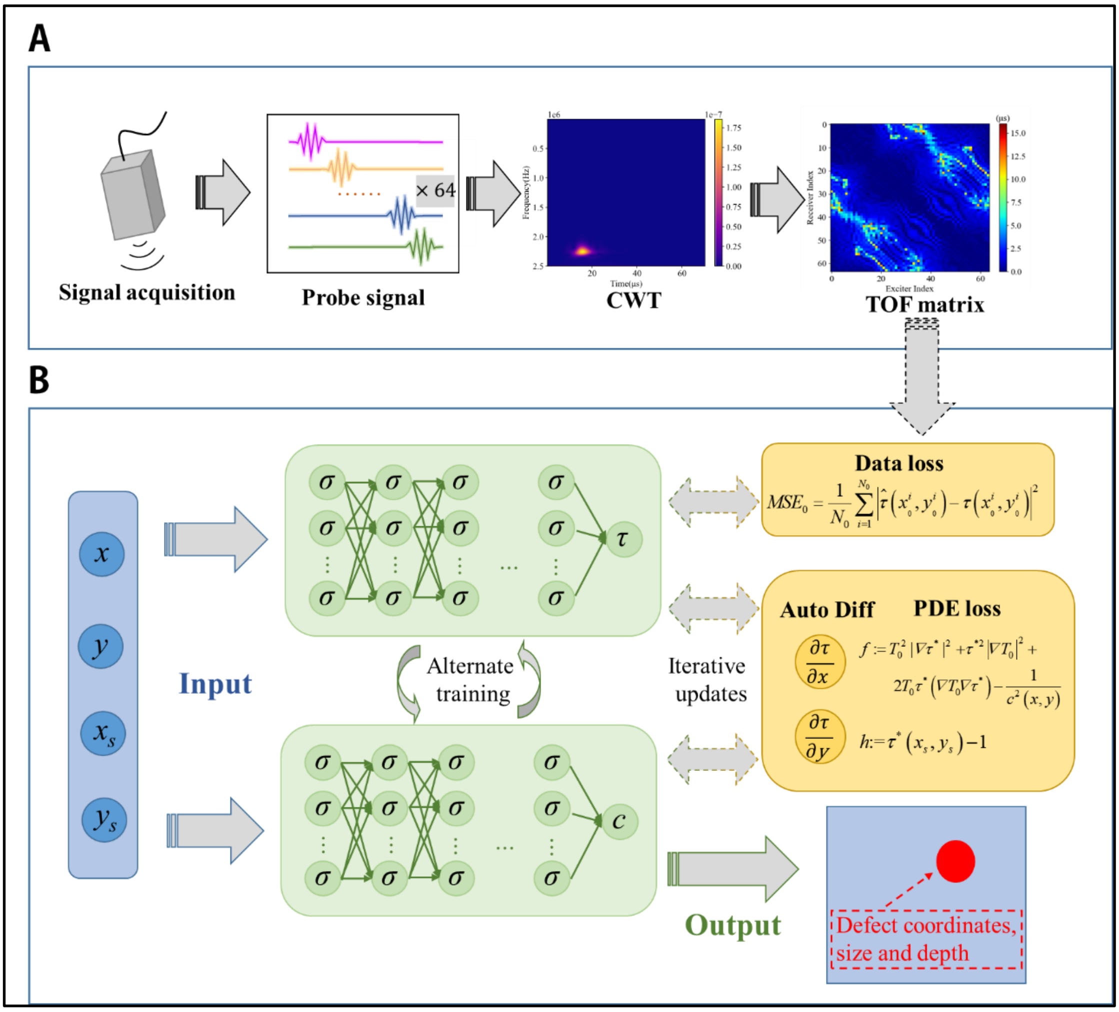


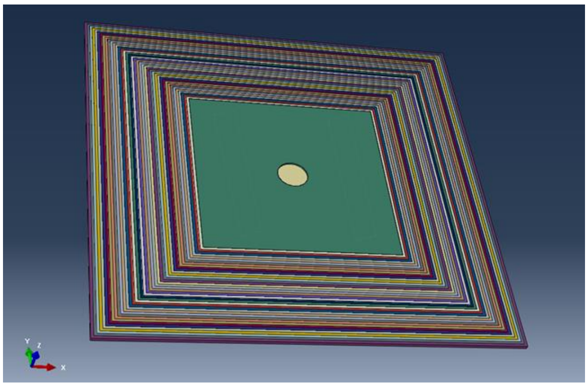
Figure 2. Overview of the PIWT framework: (a) signal acquisition and processing; (b) neural network computation framework..

3.Numerical example

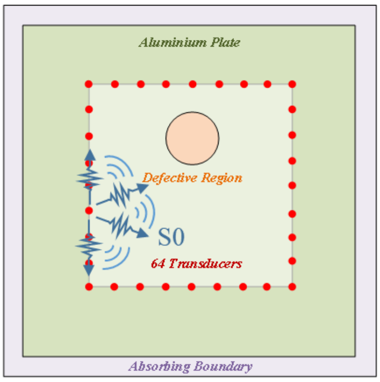
3.1. the setup of finite element model

To evaluate the efficacy of the PIWT reconstruction algorithm, a comprehensive series of simulations are conducted using ABAQUS[29]. An aluminum plate is studied and precisely measured by 150 mm x 150 mm x 3 mm, with the material properties: Young's modulus of 70.8 GPa, Poisson's ratio of 0.33, and the density of 2700 kg/m^3. In pursuit of generating an accurate depiction of the defect and capturing omnidirectional Lamb waves as they reflect off the anomaly, 64 transducers are strategically placed in a square shape, functioning alternately the emission and reception of waves. To curtail the impact of wave reflections off the boundaries of the plate, a boundary absorbing layer is meticulously modelled around the plate's perimeter, as shown in Figure.3.

A Hanning-windowed incident signal centered at 250 kHz, corresponding to a frequency-thickness product of 0.75 MHz·mm, is created to initiate a predominantly pure A_0 Lamb wave mode via one of the transducers. The wavelength for this A0 mode is determined as λ = 9.33 mm computationally. The background wave velocity of the model is set to c_0=2530 m/s, aligning with the phase velocity of the A_0 mode as depicted by the dispersion curves in Figure.1. The plate thickness is approximately 0.43λ, modelled by 10 elements along the direction of the plate thickness. A time step increment of 10^{-7}s is chosen, alongside a maximum mesh size of 1 mm, the convergence of which is rigorously achieved to ensure the precision of the calculations. To compile exhaustive data, all 64 transducers are sequentially utilized as sources of excitation in the simulation process.



(a)



(b)

Figure 3. Overview of the ABAQUS-based simulated model featuring simulation modeling (a) and

transducer distribution diagram (b).

3.2. Simulation Data Processing

In this section, the objective is to extract travel times from the displacement history as the output signals obtained from simulations. Thus, continuous wavelet transform for signal processing is employed and its mathematical definition[30] is written as:

$$W_{\psi}(a,b)=\frac{1}{\sqrt{a}}\int_{-\infty}^{+\infty}x(t)\psi^{*}\left(\frac{t-b}{a}\right)\mathrm{d}t \tag{11}$$

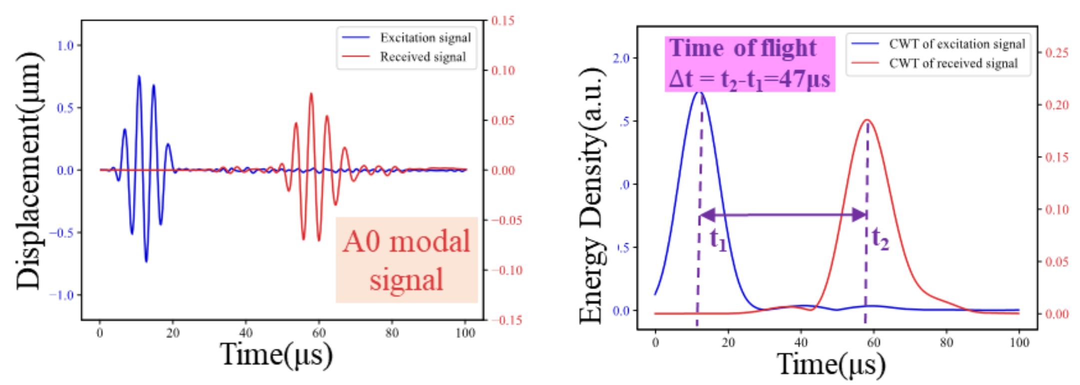
where a, b, $\psi(t)$ and $\psi^{*}(\cdot)$ are the scale parameter, time shift, basic wavelet and complex conjugate of $\psi(\cdot)$, respectively.

The energy occupied by a sampled signal can be obtained by integrating Eq. (11) as follows:

$$E=C\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty}\left|W_{\psi}(a,b)\right|^{2}\frac{da\cdot db}{a^{2}} \tag{12}$$

where the wavelet coefficients represent the energy density of the acquired signal in the time-frequency domain.

The raw excitation and reception signals, prior to any processing, are shown in Figure. 4(a). After processed by CWT, the signals are converted from the time domain to the time-frequency domain. The peak time corresponding to the energy density curve, illustrated in Figure. 4(b), is then computed. This time difference between signal excitation and reception represents the TOF of the signal.



(a) (b)

Figure 4. Signal at the 250 kHz excitation frequency showing both the excitation and received signals and the corresponding signals after CWT.

3.3 Reconstruction Results by FE simulations

To inspect the entire plate, 8 ultrasonic probes are positioned along each edge of the square detection area, resulting in a total of 1024(8*4*32) signals. The travel time signals extracted from each transducer, along with the matrix of travel time delays, are shown in Figure. 5. The travel time delay matrix represents the delays induced by defects in the plate, derived by subtracting the travel time diagram of a defect-free plate from the actual measured reception signal diagram. The results shown in Figure. 5(b) clearly highlight the time delays induced by defects, providing the information related to their shapes, positions, and depths.

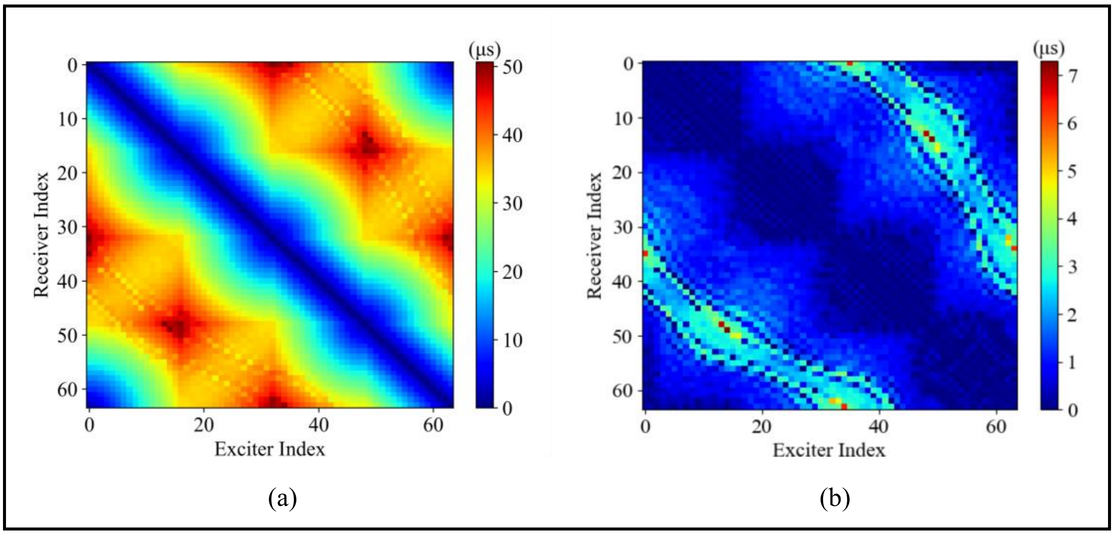


Figure 5. Single defect simulation travel time analysis featuring the reception signal matrix (a) and delay matrix (b).

The reconstruction results for a single circular defect, as well as a model containing both circular and elliptical defects, are presented in Figure. 6. These results clearly capture the contours of the defects, with the non-defect area reconstructed accurately. Regardless of the reconstruction for single or multiple defects, the defect contours are well-defined. However, slight deviations in the reconstructed shapes are observed, attributed to the limitations posed by the defect locations and transducer density. Moreover, Table 1 provides the reconstructed deepest plate thickness and the associated errors in these measurements. The data demonstrates that the defect locations are reconstructed with high precision, and the errors in thickness estimation remain within 5.5%. These improvements are primarily due to the effective use of physical information as constraints to guide the iterative process of PIWT.

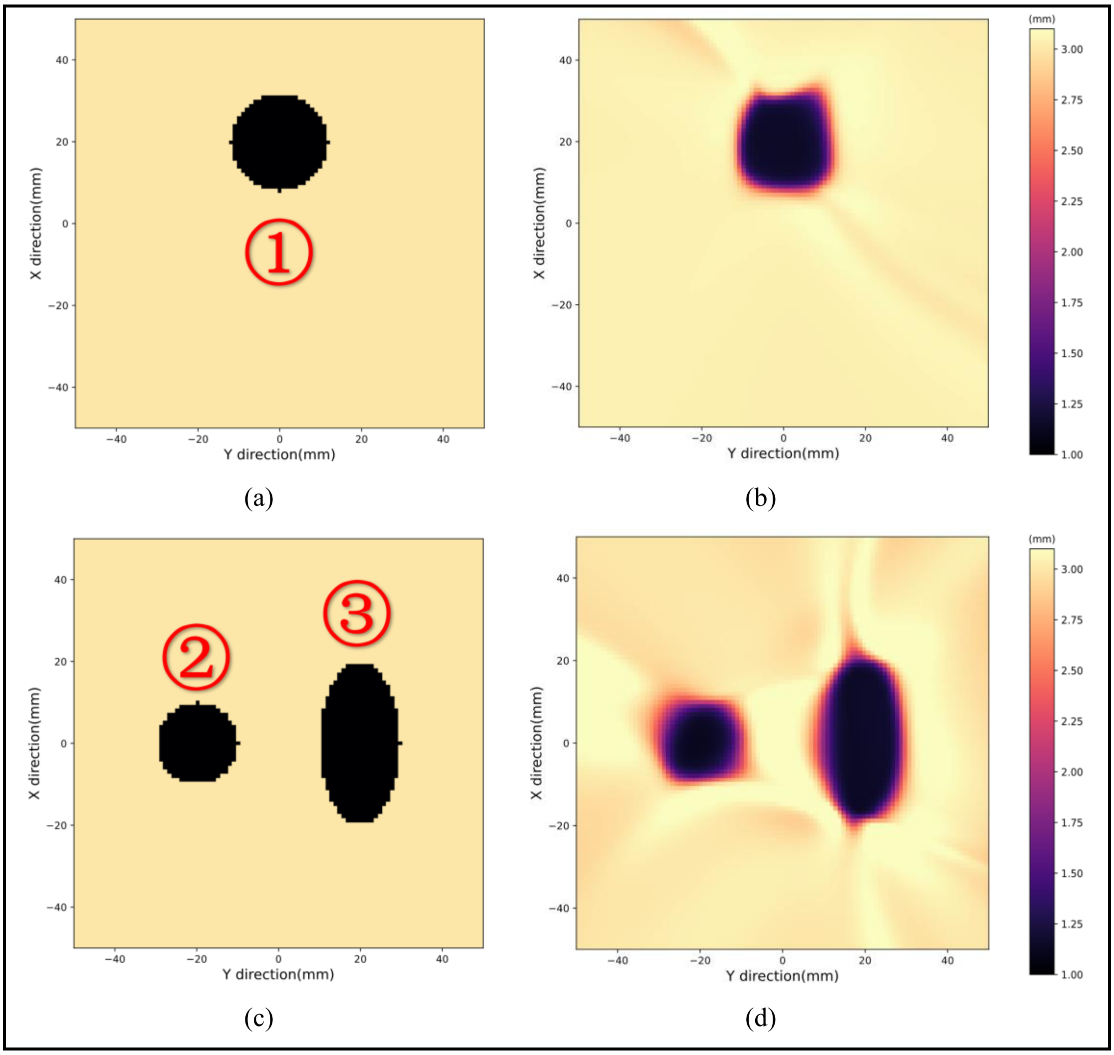


Figure 6. Reconstruction outcomes for simulated multiple circular defects: (a) and (c) the actual thickness distribution; (b) and (d) the reconstructed thickness distribution..

Table 1. Assessment results of the nominal and reconstructed defect information regarding the deepest thickness for different simulated defects.

	Num.①			Num.②			Num.③		
	Nom. (mm)	Rec. (mm)	Error (%)	Nom. (mm)	Rec. (mm)	Error (%)	Nom. (mm)	Rec. (mm)	Error (%)
Thickness	1	1.150	5	1	1.128	4.20	1	1.165	5.50
X-coordinate	0	1.065	0.3	-20	-19.393	0.61	20	20.738	0.74
Y-coordinate	20	19.79	0.2	0	-1.58	1.58	0	4.25	4.25

4.Experimental Validations

4.1 Experimental setup

To further validate the PIWT and assess the influence of the probe quantity on the reconstruction accuracy of various defects, in-lab experimental tests are conducted. Specifically, equilateral triangle and elliptical defects are fabricated in an aluminum plate, as shown in Figure. 7. Non-contact air-coupled probes (Japan Probe 0.2 K 14 × 20 N) are employed as transducers, precisely positioned in consistence with the setup used in FE simulations. A 5-cycle Hanning-windowed tone-burst signal with a central frequency of 200 kHz is used for wave generation, integrated into a non-destructive testing (NDT) system (JPR-600C) for both signal excitation and reception. To minimize the interference from the reflected waves, damping glues are applied to the edges of the aluminum plate. The air-coupled probes are suspended above the reference points, and their separation distance is systematically varied to accommodate different detection paths. To induce a relatively pure A0 mode Lamb wave, the deflection angle of the probe is determined in compliance with Snell's law. Specifically,

$$\theta_{A_0} = \sin^{-1}\left(c_{\text{air}} / c_{\text{Lamb}}\right) \tag{13}$$

where c_{air} = 340m/s is the propagation velocity of ultrasonic waves in air, while the phase velocity of the A_0 mode guided wave in the 3 mm aluminum plate at a frequency (f) of 200 kHz is c_{Lamb} = 2399 m/s. Consequently, a deflection angle θ_{A0} of approximately 10° (the computed value of 9.74°) is selected, complying with Snell's law. It is pertinent to note that the received signal comprises approximately 10 cycles, instead of the initially specified 5 cycles. This discrepancy can be attributed to two primary reasons: Firstly, wave dispersion, an inherent characteristic, distorts the signal. Also, the air-coupled transducer incorporates a built-in filter, enhancing the number of wave packets for narrowband signals. Despite these considerations, the practical impact on signal processing in the frequency domain remains minimal.

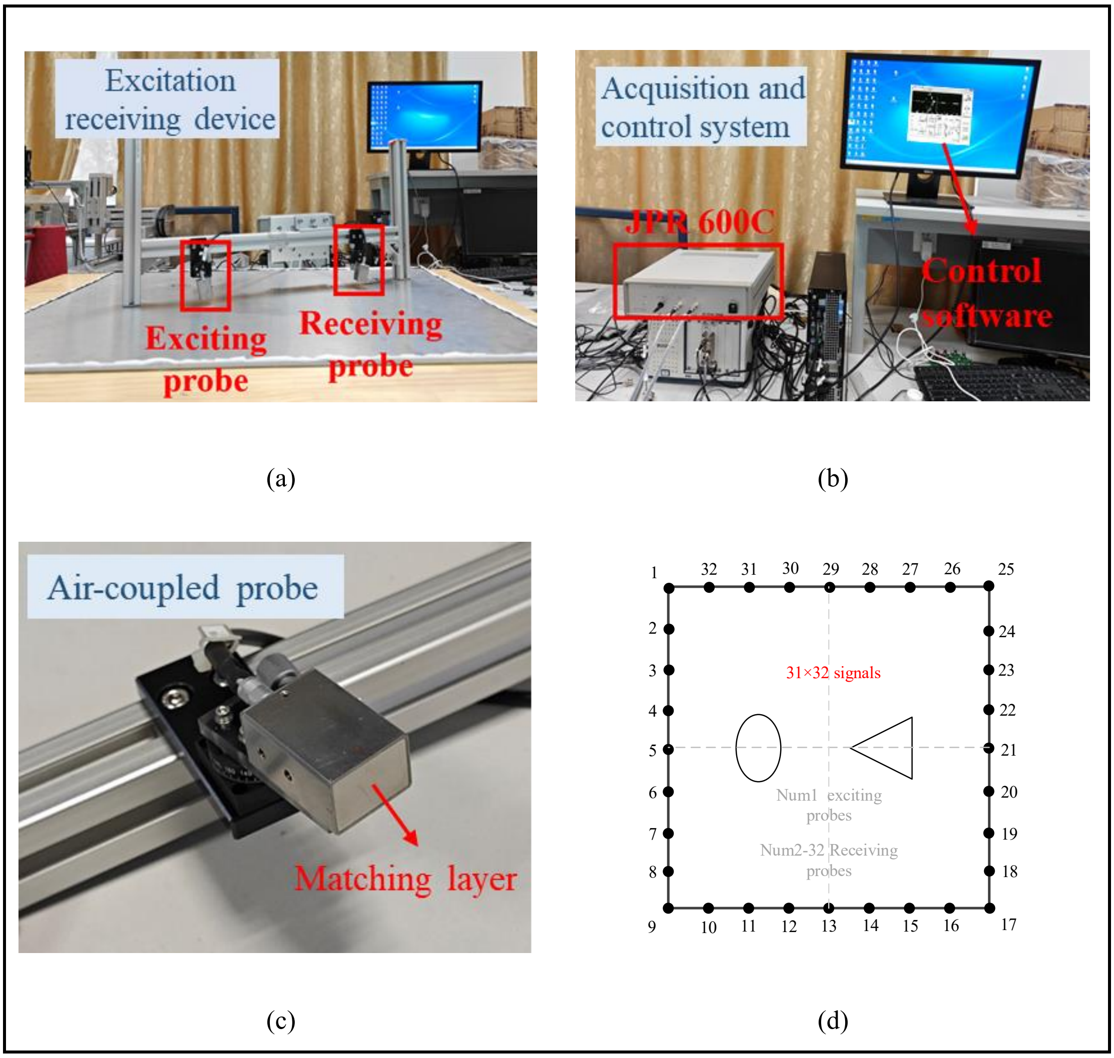


Figure 7. (a) Excitation and receiving device; (b) acquisition and control system; (c) air-coupled probe; (d) schematic diagram of exciting and receiving points.

4.2 Experimental Data Processing

The signals collected by the experimental reception probes are shown in Figure. 8(a), which simultaneously displays the signals and reflections of both S0 and A0 modes. By referencing the velocity and frequency-thickness relationship depicted in Figure. 1, it is possible to identify the wave packets corresponding to A0 and S0 modes. Further processing involves applying a continuous wavelet transform (CWT) to extract the envelope, as illustrated in Figure. 8(b). It is worth noting that, after extracting the signal travel time, the propagation time of the guided wave through the air from the air-coupled excitation source to the metal plate must be subtracted.

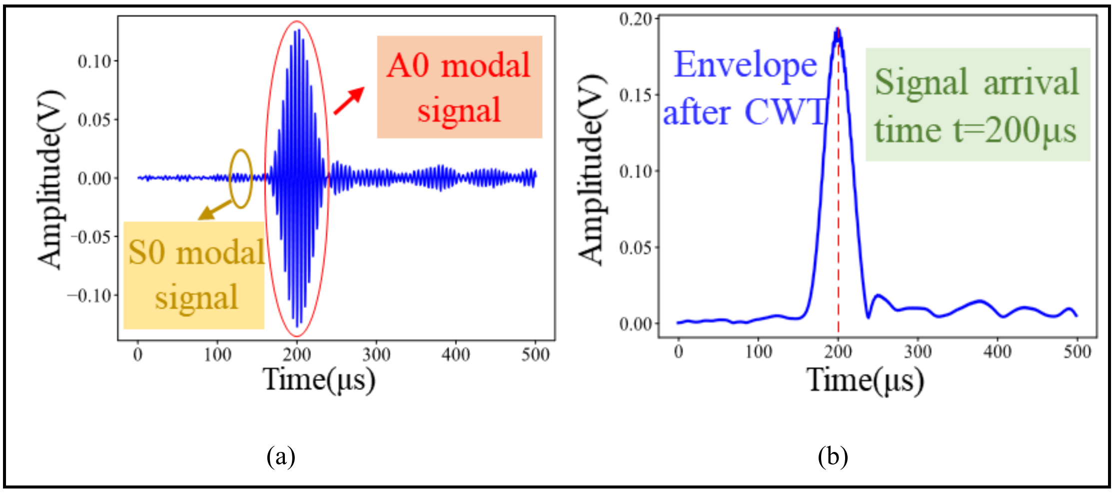


Figure 8. Experimental signal corresponding to the excitation frequency of 200 kHz: (a) the original signal; (b) the signal processed by continuous wave transformation (CWT).

4.3 Experimental results

The reconstruction results by PIWT using the experimental data are shown in Figure. 9. For the deepest regions of the defects, the positional accuracy of the elliptical defects is within 0.9%, and for the triangular defects, the error does not exceed 1.7%. Regarding the thickness reconstruction, for the 2 mm thick elliptical defects, the thinnest part of the reconstructed defects is underestimated by 1.33%. For the 1 mm thick triangular defects, the thinnest reconstructed thickness is underestimated by 15%, with this remarkable discrepancy primarily attributed to the shadowing effects of adjacent defects. Another factor influencing the reconstruction quality is the large spacing between ultrasound probes. The insufficient density of the ultrasonic array is the primary reason for shape distortion in the upper right corner of the triangular defects, resulting in the X-shaped artifacts, which cause the defects to appear blurred[31]. As the array density increases, these artifacts should be mitigated.

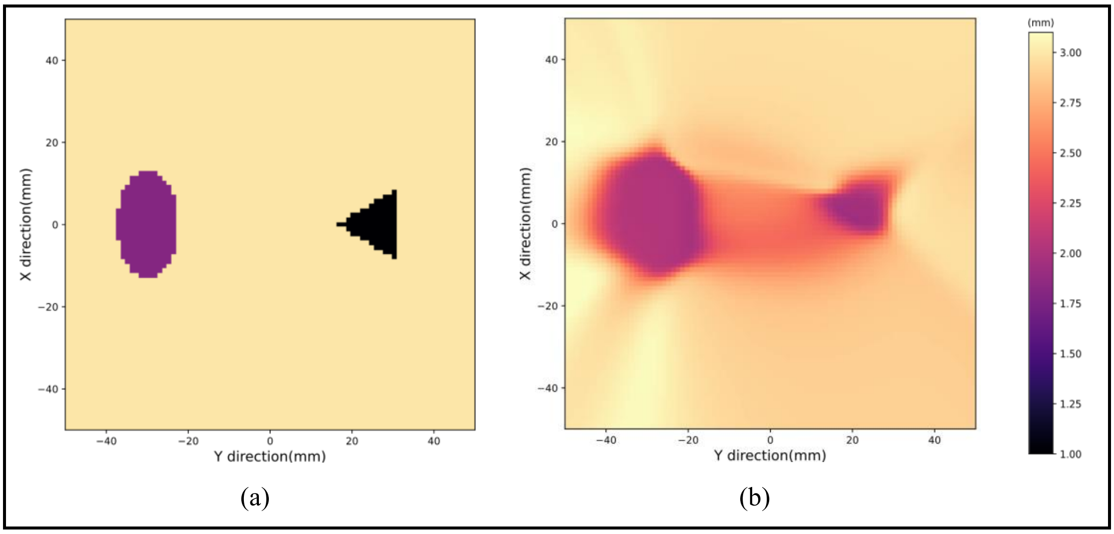


Figure 9. Schematic of the reconstruction results for dual defects in the experiment: actual thickness distribution (a) and reconstructed thickness distribution (b).

5. Data Interpretability Analysis

In this section, a visual analysis is conducted to identify the intrinsic relationship between the input and output data of the network by dimensionality reduction technique for the understanding of PIWT capability of defect reconstruction. First, this approach explores the spatial correlations between network inputs and outputs after processing data associated with various defect geometries. Then, the influence of the number of input probes on the representation in the low-dimensional space is investigated by regulating the quantity of label data fed into the PIWT. Finally, the network's capability y of modelling the complex evolution of defect/ while maintaining the physical consistency is systematically discussed.

5.1 Brief review of UMAP

Uniform Manifold Approximation and Projection (UMAP) is an efficient nonlinear dimensionality reduction algorithm widely used for the visualization and analysis of high-dimensional data. The core idea of the algorithm is to preserve the features of high-dimensional data samples in the original space while mapping them into a lower-dimensional subspace. In this process, the sample set in the high-dimensional space, denoted as $\boldsymbol{X}(X_1, X_2 \cdots X_n) \in {}^{d\times n}$, with n samples and d dimensions, is projected into a set of samples $\boldsymbol{Y} = [y_1, y_2 \cdots y_n] \in {}^{p\times n}$ in a lower-dimensional embedding space with p dimensions. This mapping is achieved by constructing a probability distribution on the high-dimensional data and then creating a corresponding probability distribution in the low-dimensional space. The principle of the mapping is to maximize the similarity between these two probability distributions, therefore realizing an effective dimensionality reduction. UMAP first calculates a conditional probability $p_{j|i}$ to represent the likelihood of the points X_i and X_j as its neighbor:

$$p_{j|i} = \begin{cases} \exp\left(-\dfrac{\| X_i - X_j \|_2 - \rho_i}{\sigma_i}\right) & X_j \in \boldsymbol{N}_i \\ 0 & \text{Otherwise} \end{cases} \tag{14}$$

where $\|\cdot\|_2$ denotes the Euclidean norm, $\boldsymbol{N}_i = \left\{X_{i,1}, X_{i,1} \cdots X_{i,k}\right\}$ represents the set of neighborhood points of X_i, and ρ_i denotes the distance from X_i to its nearest neighbors:

$$\rho_i = \min\left\{\| X_i - X_j \|_2 \mid 1 \le j \le k\right\} \tag{15}$$

Meanwhile, in equation (14), σ_i is the scale parameter, which is computed such that the total similarity between the point X_i and its *k* nearest neighbors is normalized. This is achieved by performing a binary search to find the value that satisfies:

$$\exp\left(-\frac{\| X_i - X_j \|_2 - \rho_i}{\sigma_i}\right) = \log_2(k) \tag{16}$$

As Eq.14 defines a directional similarity measure, in order to obtain a symmetric measure with respect to *i* and *j*, it needs to be symmetrized into $\boldsymbol{P}_{ij}$, which is used to represent the similarity of the data in the high-dimensional space:

$$\boldsymbol{P}_{ij} = p_{j|i} + p_{i|j} - p_{j|i} p_{i|j} \tag{17}$$

Meanwhile, in the low-dimensional embedding space, the probability of point y_i being a neighbor of point y_j can be calculated on the basis of the similarity between these two points:

$$q_{ij} = (1 + a\| y_i - y_j \|_2^{2b})^{-1} \tag{18}$$

where a and b are set to 1.929 and 0.7915, respectively.

Finally, the cost function in the UMAP algorithm is optimized using the gradient descent to obtain the corresponding samples in the low-dimensional subspace for the realization of the optimal low-dimensional data representation. The cost function shown in Eq 19 depicts a symmetric, log-transformed binary cross-entropy, where the first term represents the attraction between neighboring points in the embedding. This condition holds only when $p_{ij} \neq 0$, which means that either X_j is a neighbor of X_i, X_i is a neighbor of X_j or both. The second term represents the repulsive force, which pushes non-neighboring points further apart in the low dimensional space[32].

$$c = \sum_{i=1}^{n} \sum_{j=1, j \neq i}^{n} \left[p_{ij} \ln \frac{p_{ij}}{q_{ij}} + (1 - p_{ij}) \ln \frac{1 - p_{ij}}{1 - q_{ij}} \right] \tag{19}$$

5.2 Visualization of Inputs and Outputs

The classification results of different defects using UMAP for inputs of the TOF matrix and velocity field for are shown in Figure.10. It is noted that after dimensionality reduction, the TOF matrix data and velocity field data exhibit distinct clustering behavior for different types of defects. This indicates that the developed PIWT has captured the intrinsic distribution characteristics of the TOF data in the high-dimensional space during the training process. Throughout the latent layers of PIWT, the inputs are successfully transformed into the features represented by the output velocity field, preserving the distribution characteristics of the TOF and velocity field after dimensionality reduction. This observation elucidates how PIWT, operating as a black-box model, transforms input data within a latent high-dimensional space in accordance with intrinsic data distribution characteristics. Moreover, this intriguing finding introduces a novel strategy for the efficiency of network initialization. By mapping a TOF matrix corresponding to a specific defect type into the low-dimensional space obtained via UMAP, the defect type can be discerned based on the spatial position of one new unknown data. Subsequently, the associated velocity field is employed for the initialization of PIWT. This strategy enables the model to update its parameters in alignment with the predetermined velocity field of one specific defect type, advancing an effect analogous to the pre-training process of the velocity field for the substantial reduction of the convergence time.

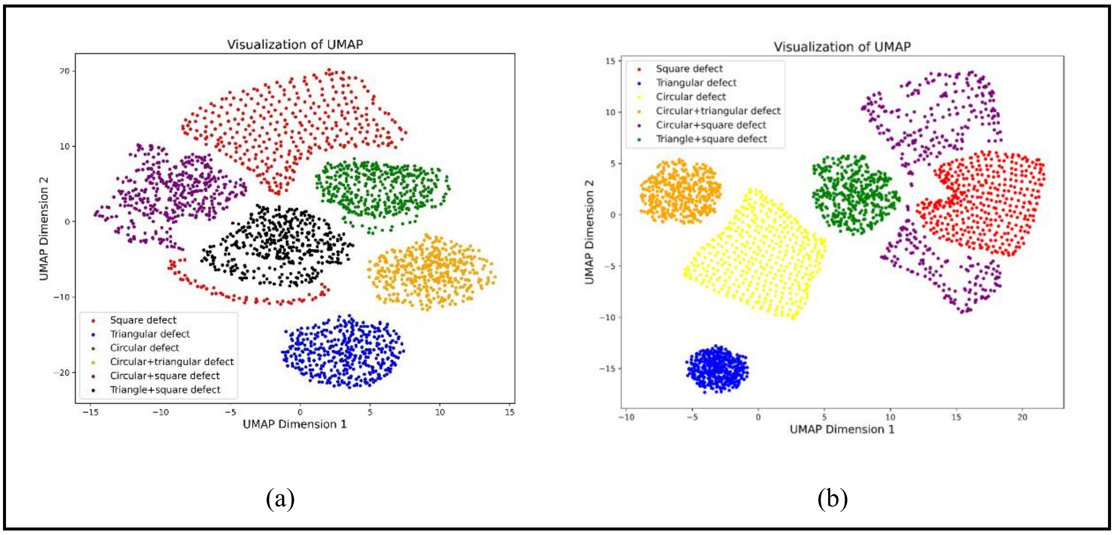


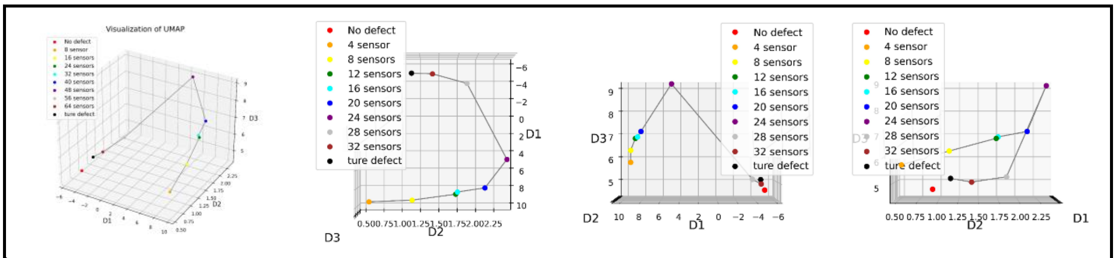
Figure 10. (a) 2D visualization of TOF matrix data for different defect types and (b) 2D visualization of

velocity field data for different defect types in the reduced-dimensional space after UMAP.

5.3 Effect of Sensor Data Volumes on the Results

Investigating the experimental tests, it is observed that the quantities of travel time data from the different numbers of sensors lead to different reconstruction outcomes. Therefore, it is essential to analyze how the volume of sensor data impacts the prediction accuracy of defect reconstructions by PIWT. In this study, data from 32 sensors are divided into eight groups, each of which is progressively added into the training dataset for the enhanced prediction model. Thus, there are eight different prediction outcomes by PIWT. Following that, these results are visualized in a 3D view using UMAP to easily explore their complex variations in high-dimensional space.

Figure 11 shows the results represented in a low-dimensional space. For circular and double ellipses defects types, it is observed that the prediction outcomes exhibit a clear trend: as the sensor data volume increases, the prediction progressively converges toward the true solution, demonstrating a well-defined trajectory. In terms of the underlying equations solving, the governing equations are embedded into the network's loss function as an extra constraint, while sensor data provides essential constraints that steer the optimization process. With the introduction of these constraints, the network's output increasingly converge to the true solution. This behavior indicates that, as the amount of data increases, the network effectively replaces the explicit governing equations with the tailored constraint physical information, thereby shedding light on how it assimilates a variety of constraints during the network learning process. Moreover, the above results offer valuable insights into its interpretability. The evolution of latent representations clearly demonstrates how the network synergizes both physical laws and sensor-derived constraints, enabling a transparent mapping from input conditions to output solutions.



(a) (b)

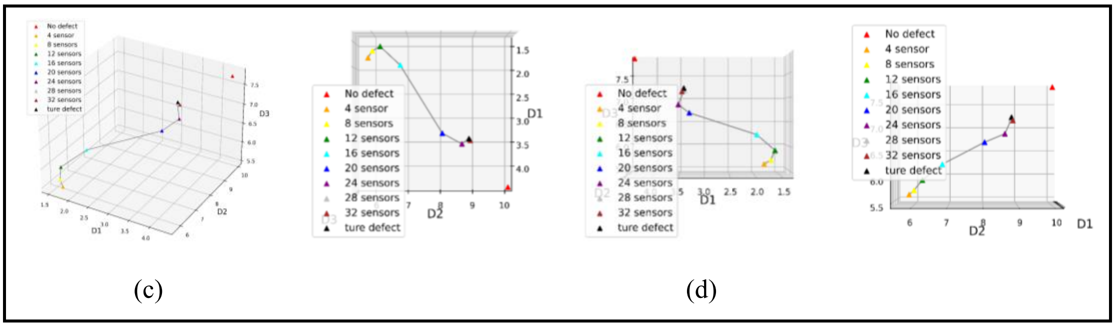


Figure 11. Spatial distribution of reconstruction results after UMAP dimensionality reduction with different sensor data quantities: (a) dimensionality reduction of reconstruction results for the defect shown in Figure 6(a); (b) frontal views of the data in (a) from three different angles; (c) dimensionality reduction of reconstruction results for the defect shown in Figure 7(a); (d) frontal views of the data in (c) from three different angles.

5. 4 Effect of Different Data from Sensor Sequences on Results

In Section 5.3, the effect of sensor data volumes as constraint data on predictions is explored, offering insights into how PIWT combines data to solve the governing equations effectively. However, the interaction between constraint data (or sensor sequences) and the governing equations within the network requires further investigation for the better interpretability of PIWT. This section discusses the effects of considering the probe data from different sequences of sensors on the prediction results of a single defect and two types of defects, respectively. Also, how the order of sensor data as an input affects the results is examined. Specifically, the influence of different constraint sequences on the prediction results are investigated when the velocity field c is fixed. Alternatively, the impact of the same constraint sequence on the results is discussed when the velocity field c varies.

Figures 12(a) and 12(b) clearly demonstrate that, for a same defect, reconstruction results using the different probe sequences as inputs initially evolve from different starting positions, and ultimately converge to the same endpoint. The overall trajectory remains consistent, which indicates that for a given velocity field, the order of constraint information provided in the network training process leads to different paths to convergence. Moreover, the change in step sizes at each iteration quantitatively reflects the amount of defect-related information as each sensor contains the defect information to some extent: sensors located closer to the defect provide more substantial information, thereby causing a more pronounced shift of the prediction toward the true solution in the low dimensional space.

On the other hand, Figures 12(c) and 12(d), which pertain to different defect types, reveal completely divergent convergence trajectories even when the same sequence of data is applied. This observation underscores that the PIWT network, during its training process, effectively learns the physical law from the governing equations for defect reconstructions, taking advantages of the distinct velocity fields. Consequently, applying an identical constraint sequence in different contexts of defect specifications results in remarkably different behaviors of convergence. Such findings not only offer a new perspective on how the sensor location and sensor sequencing affect the prediction accuracy and efficiency, but also enhance the understanding of the network's internal working mechanism throughout the interpretability.

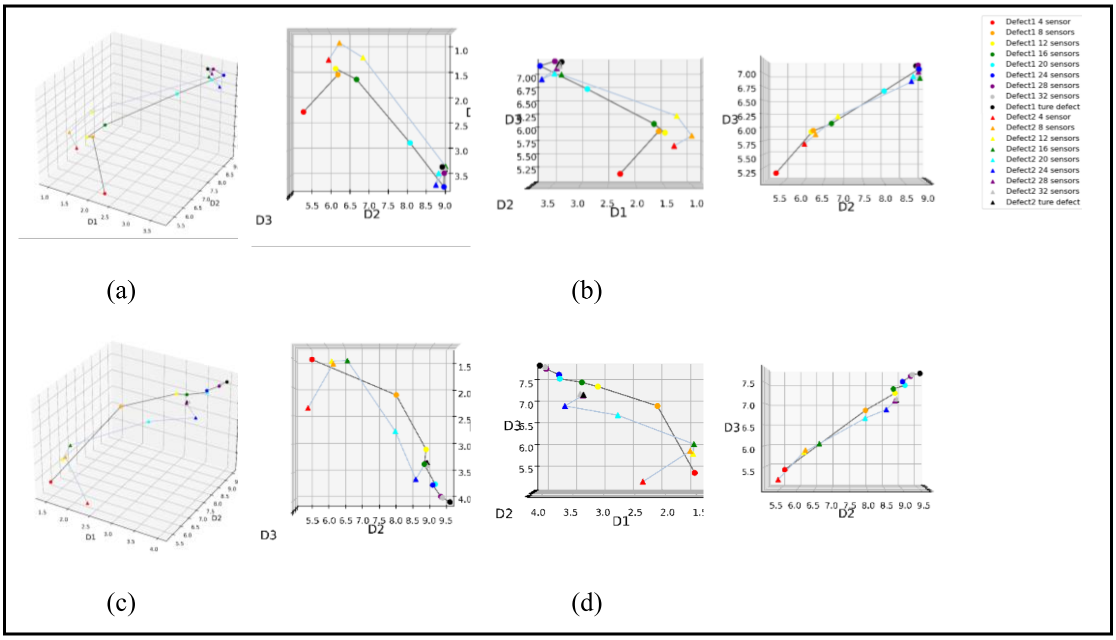


Figure.12 Impact of different data sequences on reconstruction results after UMAP dimensionality reduction: (a) dimensionality reduction of reconstruction results for the defect shown in Figure 7(a) using data with different input sequences; (b) front views of (a) from three different angles; (c) dimensionality reduction of reconstruction results for defects shown in Figure 6(a) and Figure 7(a) using data with the same input sequence; (d) front views of (c) from three different angles.

5.5. Discussion

As verified by both simulation results and experimental data, PIWT achieves the effective reconstruction of multiple defects through interpretability. This is primarily attributed to leveraging the zero-order derivative of Time-of-Flight (TOF) for constructing the initial model and incorporating the first-order derivative of TOF into the Eikonal equation to form the physical law-based loss function of the neural network PIWT. The reconstruction results by PIWT using the simulation data with 64 probes exhibit a high level of precision. From experimental data point of view, there are slight discrepancies in the vicinity of defect areas between the prediction and the ground solution. This is due to the limitation imposed by the number of probes, as only 32 ultrasound probes are deployed, resulting in a relatively large probe spacing compared to the defect size. It is noteworthy proved by Simonetti et al.**Error! Reference source not found.**, the minimum number of transducers in a circular array required to adequately sample a wavefield can be expressed as $N > 4\pi r/\lambda$, where r represents the radius of the image required to be free from grating lobes.

In the interpretability study of the PIWT model, UMAP is employed to project both the input and output data into a lower-dimensional manifold. By systematically controlling the input data, the resulting variations in the model's output can be observed and analyzed, thereby elucidating the causal relationship between the inputs and outputs. As the volume of input data increases, the outputs in the lower-dimensional space progressively converge toward the true solution, clearly demonstrating how the model incrementally guides the prediction toward the accurate ground with each additional data information.

Moreover, the study reveals that incorporating sensor data from locations in close proximity to the defect remarkably accelerates the convergence toward the true solution. This finding not only enables the rapid estimation of defect detections even when data is limited, but also provides critical insight into the model's internal mechanism. In particular, it highlights the model's ability to prioritize and effectively integrate the sensor data as constraint information, thereby enhancing the overall interpretability of PIWT. Such findings are pivotal in bridging the gap between the black-box nature of deep neural networks and the growing demand for transparent, explainable AI solutions in the areas of non-destructive testing, material science and core engineering subjects.

6. Conclusion

In this study, an improved PIWT framework is introduced for the quantitative reconstruction of multiple defects in metallic structures. By integrating prior physical knowledge into both the initial model and the loss function, the PIWT method effectively constrains the solution space and accelerates the convergence of the trunk-branch network, thus enabling a rapid preliminary reconstruction followed by refined defect localization and sizing. Simulation results—derived from two Abaqus cases—demonstrate reconstruction errors of 4.25% in position and 5.5% in depth, underscoring the framework's high accuracy. In subsequent experimental validations, despite a reduction in sensor count due to probe size limitations, the model maintained its strong generalization capability, achieving defect position errors below 1.7% and depth location errors under 15%. Moreover, the application of UMAP facilitates a clear visualization of the convergence trajectories, offering valuable insights into the latent dynamics and enhancing the interpretability of this otherwise black-box neural network. Future work will focus on introducing noise into experimental environments to further evaluate and improve the robustness of the PIWT framework for broader engineering applications.

References

[1] Groysman A. Corrosion for everybody[M]. Springer Science & Business Media, 2009.

[2] Willey C L, Simonetti F, Nagy P B, et al. Guided wave tomography of pipes with high-order helical modes[J]. Ndt & E International, 2014, 65, 8–21.

[3] Krautkrämer J, Krautkrämer H. Ultrasonic testing of materials[M]. Springer Science & Business Media, 2013.

[4] Zhao X, Royer R L, Owens S E, et al. Ultrasonic Lamb wave tomography in structural health monitoring[J]. Smart Materials and Structures, 2011, 20(10): 105002.

[5] Belanger P, Cawley P, Simonetti F. Guided wave diffraction tomography within the born approximation[J]. IEEE transactions on ultrasonics, ferroelectrics, and frequency control, 2010, 57(6): 1405-1418.

[6] Huthwaite P. Evaluation of inversion approaches for guided wave thickness mapping[J]. Proceedings of the Royal Society A: Mathematical, Physical and Engineering Sciences, 2014, 470(2166): 20140063.

[7] Williamson P R. A guide to the limits of resolution imposed by scattering in ray tomography[J]. Geophysics, 1991, 56(2): 202-207.

[8] Zhao X, Royer R L, Owens S E, et al. Ultrasonic Lamb wave tomography in structural health monitoring[J]. Smart Materials and Structures, 2011, 20(10): 105002.

[9] Dehghan-Niri E, Salamone S. A multi-helical ultrasonic imaging approach for the structural health monitoring of cylindrical structures[J]. Structural Health Monitoring, 2015, 14(1): 73-85.

[10] Yuan F G, Zargar S A, Chen Q, et al. Machine learning for structural health monitoring: challenges and opportunities[J]. Sensors and smart structures technologies for civil, mechanical, and aerospace systems 2020, 2020, 11379: 1137903.

[11] Livadiotis S, Ebrahimkhanlou A, Salamone S. A helical-based ultrasonic imaging algorithm for structural health monitoring of cylindrical structures[C]//Health Monitoring of Structural and Biological Systems XIII. SPIE, 2019, 10972: 8-14.

[12] Livadiotis S, Ebrahimkhanlou A, Salamone S. Structural health monitoring of pipelines by means of helical guided ultrasonic waves and an algebraic reconstruction technique[J]. Struct Heal Monit 2019 Proc 12Th Int Work Struct Heal Monit, 2019, 2: 1885-1892.

[13] Livadiotis S, Ebrahimkhanlou A, Salamone S. Structural health monitoring of pipelines by means of helical guided ultrasonic waves and an algebraic reconstruction technique[J]. Struct Heal Monit 2019 Proc 12Th Int Work Struct Heal Monit, 2019, 2: 1885-1892.

[14] Du M, Liu N, Hu X. Techniques for interpretable machine learning[J]. Communications of the ACM, 2019, 63(1): 68-77.

[15] Samek W. Explainable artificial intelligence: Understanding, visualizing and interpreting deep learning models[J]. arXiv preprint arXiv:1708.08296, 2017.

[16] Linardatos P, Papastefanopoulos V, Kotsiantis S. Explainable ai: A review of machine learning interpretability methods[J]. Entropy, 2020, 23(1): 18.

[17] Raissi M, Perdikaris P, Karniadakis G E. Physics-informed neural networks: A deep learning framework for solving forward and inverse problems involving nonlinear partial differential equations[J]. Journal of Computational physics, 2019, 378: 686-707.

[18] Lei N, An D, Guo Y, et al. A geometric understanding of deep learning[J]. Engineering, 2020, 6(3): 361-374. .

[19] McInnes L, Healy J, Melville J. Umap: Uniform manifold approximation and projection for dimension reduction[J]. arXiv preprint arXiv:1802.03426, 2018.

[20] Adadi A, Berrada M. Peeking inside the black-box: a survey on explainable artificial intelligence (XAI)[J]. IEEE access, 2018, 6: 52138-52160.

[21] Qian Z, Li P, Wang B, et al. A novel wave tomography method for defect reconstruction with various arrays[J]. Structural health monitoring, 2024, 23(1): 25-39.

[22] Zhou W, Li H, Yuan F G. Guided wave generation, sensing and damage detection using in-plane

shear piezoelectric wafers[J]. Smart Materials and Structures, 2013, 23(1): 015014.

[23] Rose J L. Guided wave nuances for ultrasonic nondestructive evaluation[J]. IEEE Transactions on ultrasonics, ferroelectrics, and frequency control, 2000, 47(3): 575-583.

[24] Lu Y, Michaels J E. A methodology for structural health monitoring with diffuse ultrasonic waves in the presence of temperature variations[J]. Ultrasonics, 2005, 43(9): 717-731.

[25] Qian Z, Li P, Qian Z, et al. A systematic tomography framework for thickness mapping of pipes using helical guided waves[J]. Smart Materials and Structures, 2023, 32(11): 115025.

[26] Zhang E, Dao M, Karniadakis G E, et al. Analyses of internal structures and defects in materials using physics-informed neural networks[J]. Science advances, 2022, 8(7): eabk0644.

[27] Kingma D P. Adam: A method for stochastic optimization[J]. arXiv preprint arXiv:1412.6980, 2014.

[28] Abadi M, Barham P, Chen J, et al. {TensorFlow}: a system for {Large-Scale} machine learning[C]//12th USENIX symposium on operating systems design and implementation (OSDI 16). 2016: 265-283.

[29] Qian Z, Wang B, Zhang Y, et al. Guided wave tomography of surface defects based on the method of moments[J]. Advanced theory and simulations, 2023, 6(6): 2300121.

[30] Su Z, Ye L, Bu X. A damage identification technique for CF/EP composite laminates using distributed piezoelectric transducers[J]. Composite structures, 2002, 57(1-4): 465-471. .

[31] Huthwaite P, Zwiebel A A, Simonetti F. A new regularization technique for limited-view sound-speed imaging[J]. IEEE transactions on ultrasonics, ferroelectrics, and frequency control, 2013, 60(3): 603-613.

[32] Ghojogh B, Ghodsi A, Karray F, et al. Uniform manifold approximation and projection (UMAP) and its variants: tutorial and survey[J]. arXiv preprint arXiv:2109.02508, 2021.

[33] Simonetti F, Huang L, Duric N. On the spatial sampling of wave fields with circular ring apertures[J]. Journal of Applied Physics, 2007, 101(8).