DOI 10.5162/Ultrasonic2025/P1.15

Data-Driven Frequency-Domain Full Waveform Inversion for Ultrasonic Breast Imaging

Mengting Qin[1], *Yue Zhao*[1], *and Hanlong Yin*[1]
[1] *Control Theory and Engineering, School of Astronautics, Harbin Institute of Technology,Harbin, China*
yue.zhao@hit.edu.cn

Abstract: Frequency-domain full waveform inversion is a high-resolution ultrasound method. It shows potential for breast imaging. However, it requires heavy computation and storage. These limitations restrict clinical use. We propose an end-to-end deep learning framework. This framework is constructed upon the encoder-decoder architecture of convolutional neural networks (CNNs) and is capable of directly mapping ultrasonic field data to the corresponding sound velocity distribution. Numerical simulations verify its effectiveness and efficiency.

Keywords: full waveform inversion, frequency-domain, deep learning, ultrasound tomography, breast imaging

Introduction

Ultrasound Computed Tomography (UCT) is an emerging imaging modality that employs ultrasonic transducers to emit ultrasound waves and capture interaction data with internal tissues. This enables the reconstruction of cross-sectional quantitative parameter distributions—such as sound speed and acoustic attenuation—within target tissues [1], [2]. UCT demonstrates significant potential in breast imaging by providing high-resolution images of acoustic properties within breast tissue [3].

Full Waveform Inversion (FWI), a high-resolution imaging technique originating in seismic exploration, achieves precise subsurface modeling through the utilization of full wavefield information [4]. Compared to conventional geometric reconstruction methods used in UCT, FWI establishes physics-based models grounded in wave theory equations and reconstructs tissue distributions using complete acoustic pressure data, yielding superior resolution [5]–[7]. However, FWI's high-precision imaging capability comes at the cost of substantial computational resources and storage requirements for repeated wave equation solutions, hindering its clinical adoption [8].

To mitigate computational demands, most clinical UCT implementations employ frequency-domain FWI [9], [10]. By eliminating temporal parameters and utilizing selected frequency components, frequency-domain FWI significantly reduces computational load while preserving high imaging accuracy. Nevertheless, even frequency-domain FWI necessitates iterative solutions of Helmholtz equations, resulting in prolonged computation times and high algorithmic complexity.

Recent advances in deep learning have demonstrated the capability of neural networks to approximate nonlinear operators [11], [12]. Current research focuses on leveraging neural networks to accelerate time-domain FWI reconstruction [13]–[15]. Inspired by these developments, this work proposes a deep learning-based image reconstruction method that achieves end-to-end sound speed reconstruction via frequency-domain FWI using tri-frequency breast simulation data. While conventional numerical FWI reconstruction with only three frequency bands typically fails to produce high-resolution images, our proposed network maintains satisfactory reconstruction quality while significantly reducing both input frequency requirements and computation time. Comparative studies with traditional reconstruction methods validate the effectiveness of deep neural networks in accelerating frequency-domain FWI image reconstruction.

Theory and Methods

The frequency-domain FWI imaging method is based on the frequency-domain acoustic wave equation, i.e., the Helmholtz equation. Following discretization of the imaging domain, the method establishes the functional relationship between the acoustic source s, angular frequency ω (where $\omega = 2\pi f$), velocity distribution c, and the resulting wavefield u as:

$$\left(\nabla^2 + \left(\frac{\omega}{c(\boldsymbol{r})}\right)^2\right) u(\boldsymbol{r},\omega) = -s(\boldsymbol{r},\omega) \quad (1)$$

where $\boldsymbol{r} \in \mathbb{R}^2$ is the position vector in the imaging domain, and ∇ represents the second-order spatial differential operator acting on the relevant field.

Traditional FWI imaging requires repeated solutions of the Helmholtz equation, where solving these second-order differential equations demands substantial computational resources and time. To accelerate

DOI 10.5162/Ultrasonic2025/P1.15

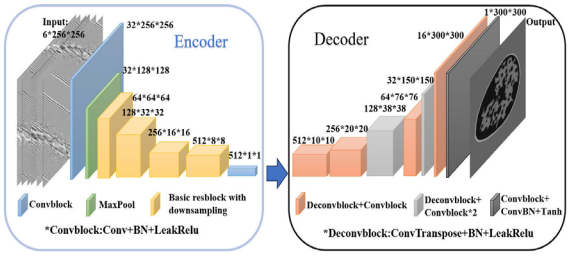

Fig. 1: The encoder-decoder framework based on CNNs.

computations, this study employs a data-driven approach to establish a direct mapping from observed wavefield data to velocity distribution.

Model: The proposed convolutional neural network (CNN) adopts an encoder-decoder architecture designed to capture complex input-output relationships, enabling direct transformation from the wavefield domain to the velocity domain (Fig. 1).

Using the encoder, high-dimensional input data undergoes feature extraction and dimensionality reduction. The encoder architecture incorporates pooling layers and diverse convolutional modules, which ultimately transform the extracted features into multichannel unit vectors. Correspondingly, the decoder receives these high-dimensional unit vectors and reconstructs the features through a deconvolution module, mapping them back to the sound velocity field.

The model input consists of tri-frequency breast wavefield observation data. Each dataset contains wavefield measurements acquired through a circumferential array of 256 equally spaced transducers functioning simultaneously as transmitters and receivers. During data acquisition, each transmitter undergoes sequential excitation while all receivers concurrently record the resulting wavefields. The output is a reconstructed sound velocity distribution map of the breast.

Loss function: The neural network's loss function integrates composite L1 and L2 norms comparing predicted and true sound velocity distributions in breast tissue, formally expressed as

$$L_\phi = L_1\left(\phi\left(u_{obs}\right) - c_{real}\right) + L_2\left(\phi\left(u_{obs}\right) - c_{real}\right) \quad (2)$$

where c_{real} denotes the true sound velocity distribution within the breast tissue and u_{obs} represents the tri-frequency wavefield measurement data. While the L2 norm constitutes the predominant objective function in FWI methodologies, it exhibits significant noise sensitivity. Conversely, the L1 norm demonstrates relative noise insensitivity. To leverage these complementary properties, this work employs a composite objective function integrating both norms.

Numerical Simulation

The breast sound velocity and wavefield observation data utilized in this study originate from the "Reconstruction of Wave Velocity in Ultrasonic Computed Tomography" competition within the AI4S Cup. The dataset comprises 7,200 paired instances of breast wavefield observations and corresponding sound velocity distributions.

The wavefield observations contain tri-frequency measurements at 0.3 MHz, 0.4 MHz, and 0.5 MHz. For each frequency, data acquisition involves sequential acoustic emission from individual transducers within a 256-element circumferential array, with all elements simultaneously functioning as receivers. Consequently, each frequency band contains 256 $\times$ 256 complex-valued measurements. To prepare these complex wavefield data for network input, real and imaginary components are separated into distinct channels. The complete breast observation dataset thus forms 6 $\times$ 256 $\times$ 256 tensors (three frequencies $\times$ two components). To mitigate near-field interference artifacts, measurements from transducers adjacent to the active emitter are nullified.

Sound velocity distributions are represented as 480 $\times$ 480 real-valued matrices. During training, only the clinically relevant central region [90:390, 90:390] (300 $\times$ 300 matrix) is retained as model output. Given the distinct characteristics of complex-valued inputs versus positive-real outputs, separate normalization schemes are applied: input data are scaled to [-1, 1], while output velocities are normalized to [0, 1]. This differentiated preprocessing is critical for optimizing imaging performance.

Numerical breast phantoms (N = 7,200) were partitioned into training (80 percent) and testing sets (20 percent). Reconstruction quality was quantified using Structural Similarity Index Measure (SSIM) and Root Mean Square Error (RMSE). Training employed 300 epochs with batch size 16, using the Adam optimizer (initial learning rate 0.001).

For traditional method comparison, benchmark FWI reconstructions were performed using competition-provided velocity phantoms. Simulations featured a 256-element transducer array (radius 3.225 cm) imaging a 480 $\times$ 480 grid (pixel size 0.15 mm). Frequency-domain FWI utilized 20 frequencies from 0.3-1.25 MHz (50 kHz increments) with 3 iterations per frequency.

Results

The quantification results of the test set subsequent to being processed by the network are presented in Tab. 1. Quantitative analysis yielded mean SSIM and RMSE

DOI 10.5162/Ultrasonic2025/P1.15

Tab. 1: Quantitative assessment results.

Quantitative Evaluation	Mean	Range
SSIM	0.88	$(0.70, 0.94)$
RMSE	9.58	$(4.60, 41.05)$

values of 0.88 and 9.58. For most test images, reconstructed SSIM exceeds 0.8 while RMSE predominantly resides within [5, 20]. Evaluation metrics reveal limited cases where predicted sound velocity distributions significantly deviate from ground truth—an inherent limitation of data-driven approaches. This method lackes explicit physical constraints and cannot achieve perfect mapping, resulting in challenging mappings from wavefield observations to velocity distributions for certain samples.

Fig. 2 comparatively demonstrates reconstruction outcomes between deep learning and conventional methods. Traditional FWI yields clearer internal tissue boundaries but suffers from circular boundary artifacts and internal artifacts caused by suboptimal sound velocity initialization or phase mismatch. Conversely, the deep learning approach produces smoother velocity distributions at the cost of some internal boundary detail. Despite localized inaccuracies, deep learning reconstructions exhibit superior image resolution.The cross-sectional sound velocity comparison reveals that the deep learning method produces a smoother velocity profile with minimal fluctuations across the continuous range, potentially limiting its ability to resolve subtle variations. Nevertheless, its overall trend closely corresponds with the true sound velocity. In contrast, the traditional inversion method can generally match the actual trend of sound velocity changes, although the reconstructed sound velocity values are not completely accurate.The root cause of this phenomenon likely stems from the network's inability to effectively incorporate frequency-related information present in the input data. This limitation hinders the network's capacity to capture the characteristic dependencies between the observed data and its frequency components. As a result, significant detail loss occurs during image reconstruction. This is particularly notable given that traditional methods typically require approximately 7-fold more observed data input than deep learning approaches.

When achieving accurate velocity mapping, the deep learning approach attains reconstruction quality comparable to conventional FWI. Notably, our method requires only three frequency observations versus twenty for traditional FWI. Crucially, the data-driven approach reconstructs breast sound velocity maps in less than 1 second versus almost 3 hours for

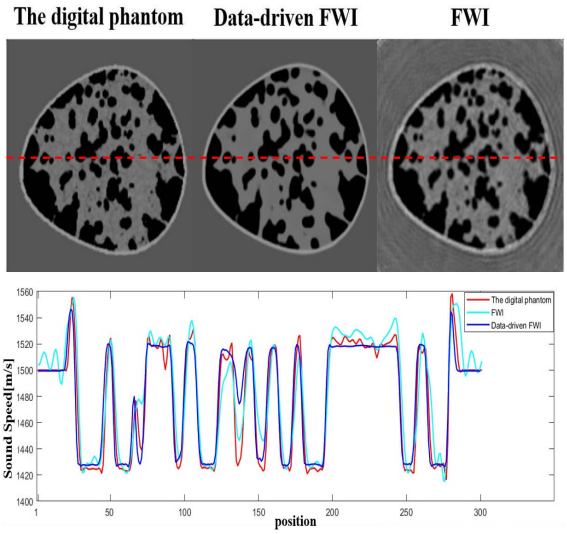

Fig. 2: Comparison of data-driven methods and traditional methods in imaging.

conventional methods. This demonstrates deep learning's potential to accelerate computationally intensive FWI techniques, suggesting trained neural networks could facilitate clinical translation of time-intensive wave velocity reconstruction.

Conclusion

This study demonstrates a data-driven FWI methodology for ultrasound breast image reconstruction. The proposed neural network employs an encoder-decoder architecture to establish direct mapping from tri-frequency wavefield observations to breast sound velocity distributions. Numerical experiments confirm the method's efficacy, with the trained network achieving high reconstruction metrics (mean SSIM: 0.88, RMSE: 9.58). Crucially, the data-driven approach generates artifact-free images comparable to conventional frequency-domain FWI results in less than 1 second versus about 3 hours. However, owing to the absence of explicit physical interpretation, deep learning methods are unable to guarantee fully accurate predictions. While deep learning reconstructions exhibit marginally reduced detail fidelity compared to traditional methods, the substantial acceleration presents a clinically viable pathway for real-time ultrasound reconstruction - particularly when balancing diagnostic utility against computational demands in clinical deployment.

ACKNOWLEDGMENT

This study was partially funded by the Natural Science Foundation of Heilongjiang province under Grant No. YQ2022F010, by the Natural Science Foundation of China (62173116, 62473108) and the postdoctoral

initial funding of Heilongjiang Province (No. LBH-Q21097).

References

[1] N. Duric, P. Littrup and C. Kuzmiak, 'Breast ultrasound tomography,' *Breast Imaging*, vol. 6, 2018.

[2] C. Li, N. Duric, P. Littrup and L. Huang, 'In vivo breast sound-speed imaging with ultrasound tomography,' *Ultrasound in medicine & biology*, vol. 35, no. 10, pp. 1615–1628, 2009.

[3] N. Duric, P. Littrup, C. Li *et al.*, 'Detection and characterization of breast masses with ultrasound tomography: Clinical results,' in *Medical Imaging 2009: Ultrasonic Imaging and Signal Processing*, SPIE, vol. 7265, 2009, pp. 457–464.

[4] A. Tarantola, 'Inversion of seismic reflection data in the acoustic approximation,' *Geophysics*, vol. 49, no. 8, pp. 1259–1266, 1984.

[5] L. Guasch, O. Calderón Agudo, M.-X. Tang, P. Nachev and M. Warner, 'Full-waveform inversion imaging of the human brain,' *NPJ digital medicine*, vol. 3, no. 1, p. 28, 2020.

[6] F. Lucka, M. Pérez-Liva, B. E. Treeby and B. T. Cox, 'High resolution 3d ultrasonic breast imaging by time-domain full waveform inversion,' *Inverse Problems*, vol. 38, no. 2, p. 025 008, 2021.

[7] N. Zhang, Y. Zhao, Y. Yuan, Y. Xiao, M. Qin and Y. Shen, 'Cross-correlation adjustment full-waveform inversion with source encoding in ultrasound computed tomography,' *Ultrasonics*, vol. 142, p. 107 392, 2024.

[8] K. Wang, T. Matthews, F. Anis, C. Li, N. Duric and M. A. Anastasio, 'Waveform inversion with source encoding for breast sound speed reconstruction in ultrasound computed tomography,' *IEEE transactions on ultrasonics, ferroelectrics, and frequency control*, vol. 62, no. 3, pp. 475–493, 2015.

[9] R. Ali, T. M. Mitcham, T. Brevett *et al.*, '2-d slicewise waveform inversion of sound speed and acoustic attenuation for ring array ultrasound tomography based on a block lu solver,' *IEEE transactions on medical imaging*, 2024.

[10] C. Zhou, K. Xu and D. Ta, 'Frequency-domain full-waveform inversion-based musculoskeletal ultrasound computed tomography,' *The Journal of the Acoustical Society of America*, vol. 154, no. 1, pp. 279–294, 2023.

[11] G. Wang, J. C. Ye and B. De Man, 'Deep learning for tomographic image reconstruction,' *Nature machine intelligence*, vol. 2, no. 12, pp. 737–748, 2020.

[12] L. Lozenski, H. Wang, F. Li *et al.*, 'Learned full waveform inversion incorporating task information for ultrasound computed tomography,' *IEEE transactions on computational imaging*, vol. 10, pp. 69–82, 2024.

[13] W. Zhang and J. Gao, 'Deep-learning full-waveform inversion using seismic migration images,' *IEEE Transactions on Geoscience and Remote Sensing*, vol. 60, pp. 1–18, 2021.

[14] Y. Wu and Y. Lin, 'Inversionnet: An efficient and accurate data-driven full waveform inversion,' *IEEE Transactions on Computational Imaging*, vol. 6, pp. 419–433, 2019.

[15] T. Robins, J. Camacho, O. C. Agudo, J. L. Herraiz and L. Guasch, 'Deep-learning-driven full-waveform inversion for ultrasound breast imaging,' *Sensors*, vol. 21, no. 13, p. 4570, 2021.