Evaluation of U-Net for Time-Domain Full Waveform Inversion Improvement

SHOAIB ANWAR, AUSTIN YUNKER, RAJ KETTIMUTHU, MARK A. ANASTASIO, UMBERTO VILLA and JIAZE HE

ABSTRACT

Ultrasound computed tomography (USCT) is one of the advanced imaging techniques used in structural health monitoring (SHM) and medical imaging due to its relatively low-cost, rapid data acquisition process. The time-domain full waveform inversion (TDFWI) method, an iterative inversion approach, has shown great promise in USCT. However, such an iterative process can be very time-consuming and computationally expensive but can be greatly accelerated by integrating an AI-based approach (e.g., convolution neural network (CNN)). Once trained, the CNN model takes low-iteration TDFWI images as input and instantaneously predicts material property distribution within the scanned region. Nevertheless, the quality of the reconstruction with the current CNN degrades with the increased complexity of material distributions. Another challenge is the availability of enough experimental data and, in some cases, even synthetic surrogate data. To alleviate these issues, this paper details a systematic study of the enhancement effect of a 2D CNN (U-Net) by improving the quality with limited training data. To achieve this, different augmentation schemes (flipping and mixing existing data) were implemented to increase the amount and complexity of the training datasets without generating a substantial number of samples. The objective was to evaluate the enhancement effect of these augmentation techniques on the performance of the U-Net model at FWI iterations. A thousand numerically built samples with acoustic material properties are used to construct multiple datasets from different FWI iterations. A parallelized, high-performance computing (HPC) based framework has been created to rapidly generate the training data. The prediction results were compared against the ground truth images using standard matrices, such as the structural similarity index measure (SSIM) and average mean square error (MSE). The results show that the increased number of samples from augmentations improves shape imaging of the complex regions even with a low iteration FWI training data.

Shoaib Anwar, Aerospace Engineering & Mechanics, The University of Alabama, Tuscaloosa, AL, USA. Austin Yunker, Data Science and Learning, Argonne National Laboratory, Lemont, IL, USA

Raj Kettimuthu, Data Science and Learning, Argonne National Laboratory, Lemont, IL, USA

Mark A. Anastasio, Department of Bioengineering, University of Illinois Urbana-Champaign, Urbana, IL, USA.

Umberto Villa, Oden Institute for Computational Engineering and Sciences, The University of Texast

INTRODUCTION

Ultrasound computed tomography (USCT) has recently become a popular method in structural health monitoring (SHM), non-destructive evaluation (NDE), and medical imaging. USCT is relatively less expensive than other computed tomography methods and is capable of acquiring data instantaneously. Full waveform inversion (FWI), an advanced inversion method that has been introduced to solve seismic wave problems, can also be applied with USCT. The benefit of employing FWI in USCT is that it can utilize all available information in the ultrasonic signals and produce high-resolution images of internal material property distributions [1]. Hence, FWI has emerged as a cutting-edge technique in SHM, NDE, and medical imaging due to its ability to accurately reconstruct material distribution [2,3].

Time-domain full waveform inversion (TDFWI) is an iterative inversion technique that uses synthetic waveform data to fit into the observed data to estimate material properties (e.g., wave speed) [4]. However, TDFWI is a computationally intensive and time-consuming method that requires high-quality data and accurate modeling of the wave propagation physics to converge with the true material distribution.

With the advancement of deep learning techniques, data-driven approaches can be used to improve ultrasound imaging including FWI. Convolutional neural networks (CNN)-based ultrasound-guided wave imaging framework was used for corrosion modeling [5]; their model accurately predicted the sizes and locations of the corrosion but was not able to capture the internal velocity and the shapes of the corrosion. FWI, alternatively, has shown great promise in identifying internal material properties [1,6].

In recent years, deep learning approaches have been explored for a variety of seismic inversion problems [7]. Network architectures such as CNN [8], recurrent neural networks (RNN) [9], and generative adversarial networks (GAN) [7] have shown great promise in solving complex seismic FWI problems. Efforts have been made to incorporate deep learning methods with USCT-based FWI to enhance the quality of reconstructed images as well as lessen the burden of the computational cost during the inference stage for medical FWI imaging. For instance, Li et al. [10] trained a standard U-Net model with low-iteration inverted images of numerical breast phantoms. However, deep learning is still in its early stage in USCT-based SHM/NDE. Our previous study [11] showed that a 2D CNN-based FWI framework could predict, with great accuracy, the material distribution while being significantly faster than the traditional FWI iterative process. However, the performance of a neural network-integrated FWI framework depends significantly on the complexity of the material distribution. Another challenge is the availability of enough data and, in some cases, even surrogate synthetic data to train the neural network (NN) models.

This paper explores the potential of improving the quality of U-Net prediction with limited FWI data to address the aforementioned issues. Multiple datasets generated from 1 to 10 iterations of the reconstructed model have been developed to train the NN model. Different augmentation strategies were implemented to add complexity and increase the size of training datasets. These strategies included flipping the sample along its axes

Austin, Austin, TX, USA
Jiaze He, Aerospace Engineering and Mechanics, The University of Alabama, Tuscaloosa, AL, USA.
* Corresponding Author: Jiaze He, Email: jhe26@eng.ua.edu.

and mixing two samples together to create a new augmented sample. High-performance computing (HPC) was used to parallelize and thus accelerate the data generation process. The remainder of the paper is arranged as follows. First, a brief introduction to TDFWI is provided. The data generation process and the description of the proposed augmentations and the U-Net architecture are then discussed in Method and Data section. The results and discussion section covers the analysis of the performance of NN model, and the conclusion contains a summary of the study results.

TIME-DOMAIN FULL WAVEFORM INVERSION (TDFWI)

TDFWI is a non-linear inversion method. It entails iteratively solving a forward problem of a time-dependent wave equation to numerically simulate wave propagation, followed by solving an inverse problem with PDE constraint optimization. This inversion process involves conducting an adjoint simulation, calculating the gradient of the misfit between observed and synthetic data, and updating model parameters [12]. The observed data can be experimentally obtained from ultrasound scanning or numerically generated (e.g., in this paper) for development purposes. The first step is to create an initial model and define sources. The wave propagation in this initial model is solved by using a 2D spectral element solver (i.e., SPECFEM2D) to create the synthetic data. This process is called forward simulation [1]. Initially, the disparity between the synthetic and observed data (i.e., misfit) is expected to be significant. The adjoint simulation is performed to update the model. This procedure involves solving the same wave equation with the time-reversed residual data as source time functions, input as sources at the receiver locations. The interaction between forward and adjoint wavefields determines the misfit function with respect to the model parameters (e.g., wave speed). This process is repeated until the misfit decreases to a specific value, at which point a high-resolution model is reconstructed. In this study, a customized version of SeisFlows, a Python-based framework, was used to implement TDFWI.

METHOD AND DATA

This study employs generating datasets from different iterations of the FWI process, applying several augmentation strategies, and training a U-Net model with these datasets. This approach enables the implementation of a two-step reconstruction process, where FWI is used to produce low-resolution inversion images in the first step, and a pre-trained U-Net model is employed to enhance the image quality in the second step. This method facilitates faster reconstruction of high-resolution images compared to the traditional TDFWI method. Details have been described in the following subsections.

Data Generation and Processing Using HPC

A water-immersed 2D domain of 12 mm $\times$ 12 mm size was considered to create the dataset. The domain consisted of 900 spectral elements with 25 control points in each element. For data acquisition, four linear arrays, each containing 101 source elements, were placed from -5 mm to 5 mm, parallel to each boundary with a distance of 2 mm

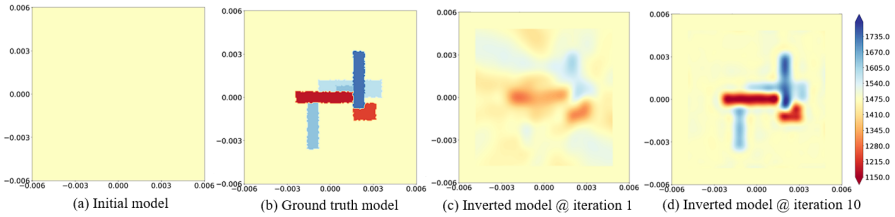

Figure 1. Longitudinal wave speed (V_p m/s) mapping of (a) initial model, (b) ground truth model and reconstructed FWI model after (c) 1st iteration and (d) 10th iteration.

(See Figure 1 in [11]). In one excitation event, all sources in a single array were simultaneously excited with an ultrasound signal (Ricker signal) of 1 MHz center frequency. As a result, four forward simulations were carried out to implement TDFWI in a single model. For the initial model, only a water environment with 1479.7 m/s wave speed was considered (Figure 1 (a)). Alternatively, the ground truth models incorporated unknown material inside the water environment (Figure 1 (b)). Multiple ground truth models were created by randomly varying the size, locations, and wave speed (ranging from 1100 to 1800 m/s) of these unknown materials. In our previous study [11], a 2D CNN was employed to improve the reconstruction quality of the model using only a single FWI iteration. However, the performance of the 2D CNN deteriorated as the complexity of the material distribution was increased by adding the number of unknown materials and widening the wave speed range. This issue could be addressed by increasing the training dataset and enhancing the quality of the data. To improve the quality of the data, the training datasets were constructed with the inverted model from higher FWI iterations. Therefore, multiple datasets from different iterations (ranging from 1-10) were created. Each dataset consisted of 1000 FWI inverted models. Figures 1 (c) and (d) show the reconstructed models at the first and 10th iterations of TDFWI. Data generation was done using Simple Linux Utility for Resource Management (SLURM). A batch of 10 samples performed 10 FWI iterations per sample using 40 CPUs in a single SLURM job. Each job took approximately 15 minutes to complete. Thus, to generate the datasets, 100 SLURM jobs were executed.

The output of each iteration was a one-dimensional array containing the wave speed value of each spectral element control point. The data processing started with the removal of the duplicate control points of neighboring spectral elements. This reduced 1D array of a shape of 1×9025 was then mapped into a 2D array of a shape of 95×95 [11]. After that, all 2D arrays of 1000 inverted models were accumulated into one big array of $1000 \times 95 \times 95$. For easy pooling and scaling, the array was padded on the right and bottom sides making the final shape of $1000 \times 96 \times 96$. As a common practice, the wave speeds were normalized between 0 and 1 using a linear regression normalization technique. Each sample is processed as a greyscale picture by the deep learning model, yielding a channel of one (instead of three for RGB images). Consequently, the final shape of the datasets was $1000 \times 96 \times 96 \times 1$. Finally, the dataset was shuffled and split into training, validation, and testing at a ratio of 80-10-10%, respectively.

Augmentations

This study employed two augmentation strategies to increase the quantity and com-

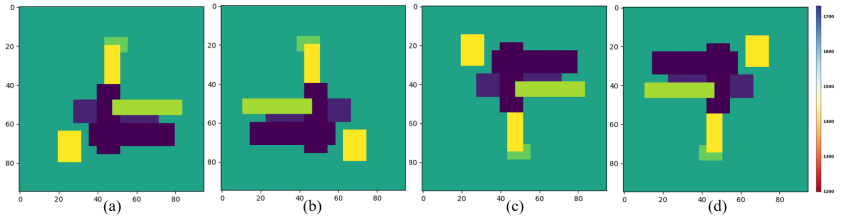

Figure 2. An example of flipping augmentation. (a) was the original sample which was flipped vertically and created a new sample (b). Then both (a) and (b) were flipped horizontally and created new samples (c) and (d).

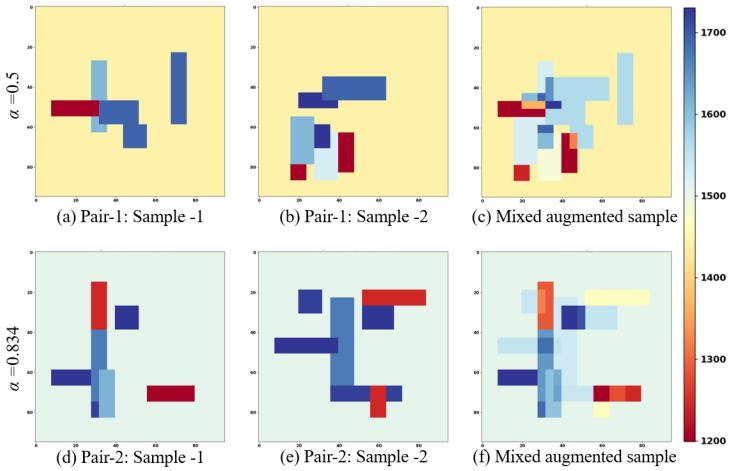

Figure 3. An example of the mixing augmentation. Two ground truth models (a) and (b) were mixed together with $\alpha = 0.5$ and generated a new sample (c). Similarly, (e) and (f) samples were mixed with $\alpha = 0.834$ and generated a new sample (f).

plexity of the training dataset without generating samples. The first strategy used was the flipping augmentation whereby the sample was flipped along its axes. This allowed us to increase the original dataset with 2400 additional training samples. Figure 2 shows an example of the flipping augmentation, where the original sample was flipped vertically to create a new sample (Figure 2 (b)) and both were flipped horizontally to generate two new samples (Figures 2 (c) and (d)). The second strategy used was mixing augmentation. Here, two samples were multiplied by coefficients α and $(1 - \alpha)$, respectively, and then added together making a new sample. The value of α randomly varied from 0 to 1. The number of augmented samples was generated to equal the number of non-augmented samples. Therefore, 1600 samples, 800 original and 800 mixed, were used for training. Figure 3 illustrates the mixing augmentation strategy, where two pairs of samples – Figures 3 (a) & (b) and (d) & (e) were mixed together and produced two new samples – Figures 3 (c) and (f), respectively – and the values of α were 0.5 in the first pair and 0.834 in the second pair.

Architecture

The task of accelerating the FWI iteration process can be posed as that of translating a low-iteration FWI image—ranging from 1-10 here—into its corresponding ground truth image. To do this, we employed the U-Net architecture (shown in Figure 4), which

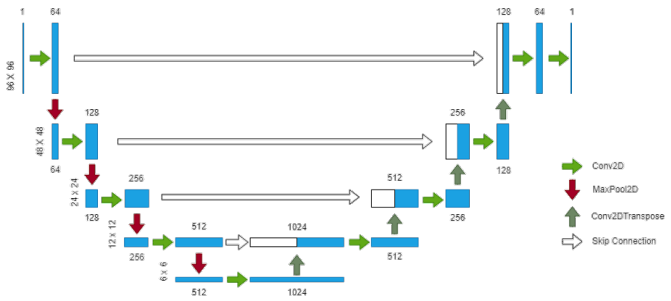

Figure 4. The architecture of the U-Net model used in this study [11].

was originally introduced for biomedical image segmentation by [13]. We followed the general U-Net architecture which is composed of four down-sampling layers (red arrows) and four up-sampling layers (dark green arrows). In the down-sampling process, four sets of two convolutional kernels extract feature maps. Then, followed by a pooling layer (light green arrows), the feature map projections were distilled to the most essential elements using a signal-maximizing process. By the end of the down-sampling process, the feature maps were 1/16 of the original size: 6x6 in Figure 4. Successful training should result in 6 channels in this feature map, which is verified by the model predictions in Figure 5. Bilinear interpolation was used in the up-sampling process to expand the feature maps. At each layer, high-resolution features from the down-sampling path were concatenated to the up-sampled output from the layer below to form a larger number of feature channels (white arrows). This structure allowed the network to propagate context information to higher-resolution layers so that the following convolution layer could learn to assemble a more precise output based on this information. Finally, the mean squared error loss function was employed in the network.

RESULTS AND DISCUSSION

The U-Net model was trained using each dataset with three variants according to the augmentation strategy: no augmentations, flipping augmentations, and mixing augmentations. The predicted images from the U-Net models were compared with the ground truth images using SSIM and MSE values.

Predictions of all U-Net models achieved SSIM values above 94%. The SSIM values increased for U-Net models trained with higher iteration inverted images, which was expected due to the higher quality of the data. When trained with the 10th iteration dataset, the model achieved more than 99% similarity to the ground truth model. Consequently, the MSE value decreased significantly at higher iteration prediction results.

Comparing the different augmentation strategies, the performance of the U-Net models was comparable in the flipping augmentations and no-augmentations for every iteration. Figure 5 illustrates the model performance for a particular test sample (Figure 5 (a)). Figures 5 (b) and (f) are the inverted images for the 2nd and 10th iterations. For lower iteration, the model struggled to reconstruct the area where different wave speed regions overlapped (Figures 5 (c), (d), and (e)). The SSIM and MSE values were very close for no-augmentation and the flipping augmentation strategies. However, the boundary of the different wave speed regions was defined better in the mixing augmenta-

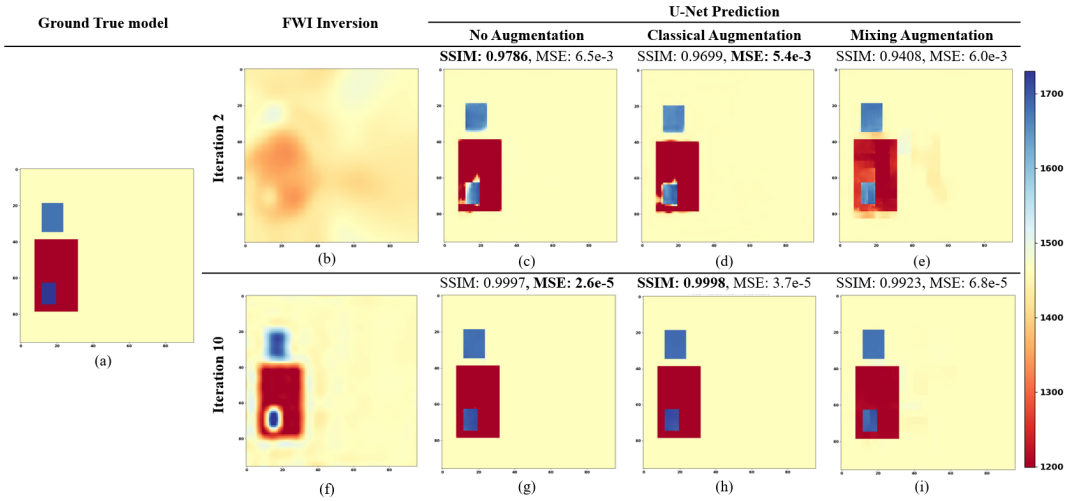

Figure 5. An example of the prediction to compare the U-Net performance. (a) is the ground truth model, (b) and (f) are the 2nd and 10th inversion images, respectively. (c), (d) and (e) are the prediction of U-Net when the model was trained with the dataset of 2nd iteration FWI samples with no augmentation, flipping augmentation, and mixing augmentation, respectively. (g), (h) and (i) are the prediction of U-Net when the model was trained with the dataset of 10th iteration FWI samples with no augmentation, flipping augmentation, and mixing augmentation, respectively.

tion than in the other two cases, though it introduced the most artifacts (different shades in yellow and red regions in Figure 5(e)). At the higher iteration, the wave speed reconstruction as well as boundary definition for different regions were exceedingly better in no-augmentations (Figure 5 (g)) and flipping augmentation (Figure 5 (h)) than the mixing augmentation (Figure 5 (i)).

The underperformance of the mixing strategy may be attributed to the higher complexity of the training dataset relative to the testing samples. For validation and testing, no augmented sample was used; this may shift the material distribution in the training dataset from the validation and test datasets. Additionally, the size of the training dataset may have played a role in this regard. Therefore, a larger dataset size with different augmentation strategies will be considered in future studies.

CONCLUDING REMARKS

This study explored the potential of improving the quality of a U-Net model prediction for limited FWI data. The model was trained with multiple datasets generated from different FWI iterations. To add complexity and increase the training dataset size, two augmentation strategies–flipping and mixing augmentation–were introduced. The results demonstrate that the model was able to reconstruct the samples even in lower iterations. At higher iterations, augmentations have no significant effects on prediction for higher iteration training data because the no-augmentation model has already reconstruct the sample almost perfectly. On the other hand, augmentations helped to improve the boundary region in prediction with lower iteration training data. Future studies include exploring different augmentation strategies and training with a larger dataset for lower iteration FWI.

ACKNOWLEDGEMENT

We appreciate the financial support by NSF projects #2152764 and #2152765, the computational resources provided by ACCESS Project #MDE220005, and the high-performance computing technical support for data generation/training by Dr. Paul Rodriguez through the XSEDE Extended Collaborative Support Service (ECSS) grant.

REFERENCES

1. He, J., J. Rao, J. D. Fleming, H. N. Gharti, L. T. Nguyen, and G. Morrison. 2021. "Numerical ultrasonic full waveform inversion (FWI) for complex structures in coupled 2D solid/fluid media," *Smart Materials and Structures*, 30(8):085044.
2. Nguyen, L. T. and R. T. Modrak. 2018. "Ultrasonic wavefield inversion and migration in complex heterogeneous structures: 2D numerical imaging and nondestructive testing experiments," *Ultrasonics*, 82:357–370.
3. Guasch, L., O. Calderón Agudo, M.-X. Tang, P. Nachev, and M. Warner. 2020. "Full-waveform inversion imaging of the human brain," *NPJ digital medicine*, 3(1):28.
4. Bleibinhaus, F. 2016. *Full-Waveform Inversion of Controlled-Source Seismic Data*, Springer Berlin Heidelberg, Berlin, Heidelberg, ISBN 978-3-642-36197-5, pp. 1–13, doi:10.1007/978-3-642-36197-5_376-1.
5. Wang, X., M. Lin, J. Li, J. Tong, X. Huang, L. Liang, Z. Fan, and Y. Liu. 2022. "Ultrasonic guided wave imaging with deep learning: Applications in corrosion mapping," *Mechanical Systems and Signal Processing*, 169:108761.
6. Li, R., H. Chen, Y. Peng, J. Li, and Y. Li. 2020. "Ultrasound computed tomography of knee joint," *Chinese Journal of Electronics*, 29(4):705–716.
7. Deng, C., S. Feng, H. Wang, X. Zhang, P. Jin, Y. Feng, Q. Zeng, Y. Chen, and Y. Lin. 2021. "OpenFWI: Large-Scale Multi-Structural Benchmark Datasets for Seismic Full Waveform Inversion," *arXiv preprint arXiv:2111.02926*.
8. Park, M. J. and M. D. Sacchi. 2020. "Automatic velocity analysis using convolutional neural network and transfer learning," *Geophysics*, 85(1):V33–V43.
9. Wang, W., G. A. McMechan, and J. Ma. 2021. "Elastic isotropic and anisotropic full-waveform inversions using automatic differentiation for gradient calculations in a framework of recurrent neural networks," *GEOPHYSICS*, 86(6):R795–R810, doi:10.1190/geo2020-0542.1.
10. Li, F., U. Villa, S. Park, and M. A. Anastasio. 2021. "3-D stochastic numerical breast phantoms for enabling virtual imaging trials of ultrasound computed tomography," *IEEE transactions on ultrasonics, ferroelectrics, and frequency control*, 69(1):135–146.
11. Kleman, C., S. Anwar, Z. Liu, J. Gong, X. Zhu, A. Yunker, R. Kettimuthu, and J. He. 2023. "Full Waveform Inversion-Based Ultrasound Computed Tomography Acceleration Using 2D Convolutional Neural Networks," *Journal of Nondestructive Evaluation, Diagnostics and Prognostics of Engineering Systems*:1–13.
12. Tromp, J., C. Tape, and Q. Liu. 2005. "Seismic tomography, adjoint methods, time reversal and banana-doughnut kernels," *Geophysical Journal International*, 160(1):195–216.
13. Shan, H., Y. Zhang, Q. Yang, U. Kruger, M. K. Kalra, L. Sun, W. Cong, and G. Wang. 2018. "3-D convolutional encoder-decoder network for low-dose CT via transfer learning from a 2-D trained network," *IEEE transactions on medical imaging*, 37(6):1522–1534.