Massive Ultrasonic Data Compression Using Wavelet Packet Transformation Optimized by Convolutional Autoencoders

Boyang Wang, *Member, IEEE*, and Jafar Saniie, *Life Fellow, IEEE*

***Abstract*— Ultrasonic signal acquisition platforms generate considerable amounts of data to be stored and processed, especially when multichannel scanning or beamforming is employed. Reducing the mass storage and allowing high-speed data transmissions necessitate the compression of ultrasonic data into a representation with fewer bits. High compression accuracy is crucial in many applications, such as ultrasonic medical imaging and nondestructive testing (NDT). In this study, we present learning models for massive ultrasonic data compression on the order of megabytes. A common and highly efficient compression method for ultrasonic data is signal decomposition and subband elimination using wavelet packet transformation (WPT). We designed an algorithm for finding the wavelet kernel that provides maximum energy compaction and the optimal subband decomposition tree structure for a given ultrasonic signal. Furthermore, the WPT convolutional autoencoder (WPTCAE) compression algorithm is proposed based on the WPT compression tree structure and the use of machine learning for estimating the optimal kernel. To further improve the compression accuracy, an autoencoder (AE) is incorporated into the WPTCAE model to build a hybrid model. The performance of the WPTCAE compression model is examined and benchmarked against other compression algorithms using ultrasonic radio frequency (RF) datasets acquired in NDT and medical imaging applications. The experimental results clearly show that the WPTCAE compression model provides improved compression ratios while maintaining high signal fidelity. The proposed learning models can achieve a compression accuracy of 98% by using only 6% of the original data.**

***Index Terms*— Convolutional autoencoder (CAE), data compression, ultrasonic signals, unsupervised learning, wavelet packet transformation (WPT).**

I. Introduction

WAVELET packet transformation (WPT) is a widely used algorithm for signal compression, denoising, and watermarking [1]–[3]. WPT is particularly efficient for compressing transient signals. Ultrasonic nondestructive testing (NDT) data and medical imaging datasets are composed of wavelets and are well suited for WPT compression [4], [5]. The autoencoder (AE) is an unsupervised learning model that can be used for dimension reduction, feature extraction, signal compression, and denoising [6]–[8]. The convolutional AE (CAE) contains convolution layers, which can achieve better performance with fewer trainable weights, especially when high-resolution images are processed [9], [10]. In this study, we designed a WPT CAE (WPTCAE) for ultrasonic signal compression. This approach offers a substantial performance enhancement compared with WPT-based signal compression. In other words, WPTCAE is a modified WPT filter bank for optimal performance determined by ultrasonic training data.

The WPTCAE model for data compression demands that the wavelet kernels have a reasonable resemblance to the signals in the dataset, as is the case for ultrasonic signals consisting of wavelets. Fig. 1 shows the design flow of the WPTCAE compression system composed of four functional blocks: 1) ultrasonic data acquisition; 2) WPT compression optimization; 3) WPTCAE compression model implementation and training; and 4) WPTCAE compression and reconstruction.

In ultrasonic NDT or medical imaging applications, the pulser/receiver is used to acquire ultrasonic A-Scans (amplitude scans), as shown in Fig. 1(A2). By scanning or beamforming, one obtains a dataset of ultrasonic B-Scans that can reveal the inner structure of materials or tissues, as shown in Fig. 1(A3). Ultrasonic B-Scans are 2-D ultrasonic images that are composed of a spatial sequence of A-Scans. Consequently, A-Scans are ideal signals for analyzing the performance of the proposed WPTCAE compression algorithm.

Fig. 1(B) shows the procedure for finding the optimal WPT compression scheme for a given ultrasonic dataset. The success of the WPT compression algorithm strongly relies on selecting the proper wavelet kernel and decomposition tree structure. To find the optimal WPT compression solution for a given dataset, a Python script is implemented to search for the optimal wavelet kernel and the structure of the WPT decomposition tree. Fig. 1(B3) shows the processing blocks involved in decomposition diagrams of the WPT compression algorithm where three subband components are selected as the compressed data. Fig. 1(B4) shows the reconstruction diagram associated with Fig. 1(B3).

Fig. 1(C) shows the procedure for building and training the WPTCAE compression model using TensorFlow [11]. The WPTCAE model has two functional blocks: an encoder for signal compression and a decoder for signal reconstruction. As shown in Fig. 1(C2), the structure and initial coefficients are designed according to the optimal WPT compression algorithm acquired in Fig. 1(B). The WPTCAE encoder/decoder is trained with the backpropagation algorithm.

Manuscript received October 27, 2020; revised May 4, 2021 and August 2, 2021; accepted August 12, 2021. *(Corresponding author: Jafar Saniie.)*

The authors are with the Department of Electrical and Computer Engineering, Illinois Institute of Technology, Chicago, IL 60616 USA (e-mail: saniie@iit.edu).

Color versions of one or more figures in this article are available at https://doi.org/10.1109/TNNLS.2021.3105367.

Digital Object Identifier 10.1109/TNNLS.2021.3105367

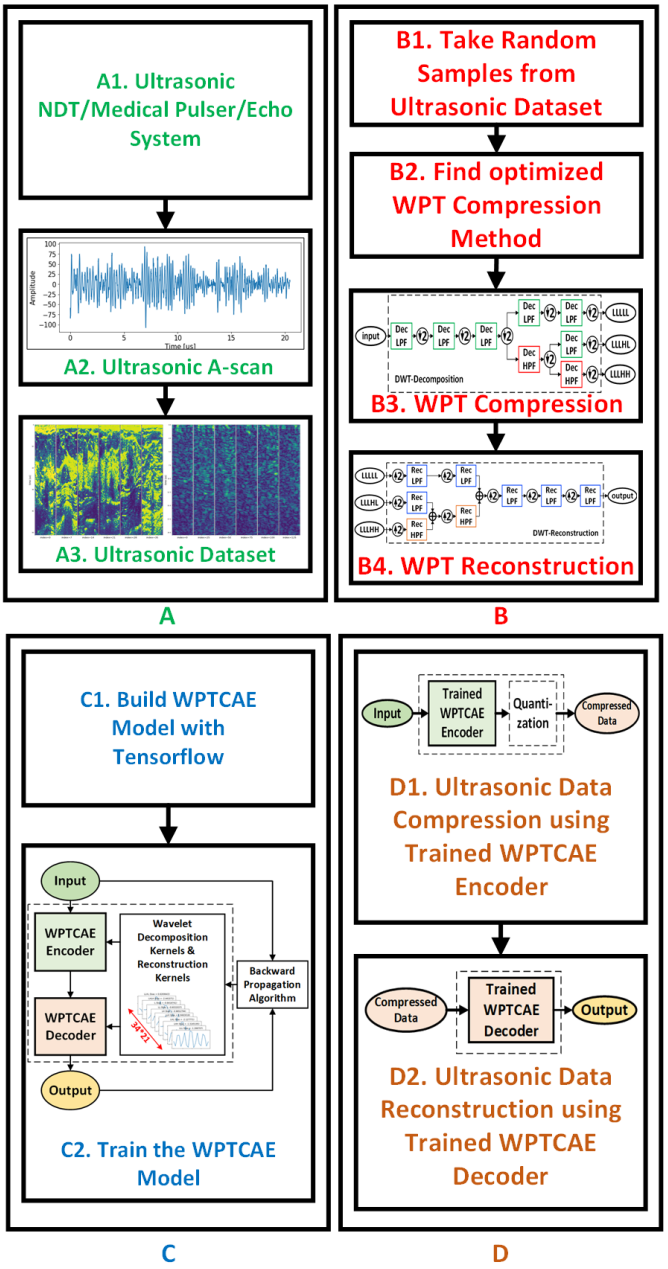


Fig. 1. Design flow of WPT optimized by a CAE for ultrasonic data compression. (A) Ultrasonic data acquisition. (B) WPT compression optimization. (C) WPTCAE model. (D) WPTCAE compression and reconstruction.

Fig. 1(D) illustrates the ultrasonic data compression and reconstruction with the trained WPTCAE model. As shown in Fig. 1(D1), the compression of the dataset is achieved by passing the ultrasonic dataset through the trained WPTCAE encoder. Fig. 1(Dg2) shows the ultrasonic data reconstruction process using the WPTCAE decoder.

The evaluation criteria for benchmarking the compression algorithms include the compression ratio (defined as the original data size over the compressed data size) and the percentile mean square error (PMSE) in the recovery of the compressed signal or Pearson's correlation coefficient (PCC). The PMSE of the reconstructed signal S_{re} and the original signal S_{org} is defined as

$$\text{PMSE}(S_{\text{org}}, S_{\text{re}}) = 100 \times \frac{\sum_{i=0}^{N}(S_{\text{org}}[i] - S_{\text{re}}[i])^2}{\sum_{i=0}^{N}(S_{\text{org}}[i])^2}. \quad (1)$$

A lower PMSE means a closer similarity in mean square sense. The PCC is defined as

$$\text{PCC}(S_{\text{org}}, S_{\text{re}}) = \frac{\text{cov}(S_{\text{org}}, S_{\text{re}})}{\sigma(S_{\text{org}})\sigma(S_{\text{re}})} \quad (2)$$

which can measure the similarity between S_{org} and S_{re}. In (2), the symbols *cov* and σ represent the covariance and the standard deviation, respectively. The value of the PCC lies between -1 and $+1$. An ideal compression algorithm has a PCC value close to $+1$. The PCC is well suited for measuring the overall similarity of the reconstructed and original signals.

To benchmark the proposed WPTCAE compression model, we used two types of ultrasonic image datasets, as described in Section II. One type is the ultrasonic NDT data that we acquired in our research laboratory [12], and the other is the Open Access Series of Breast Ultrasonic Data (OASBUD) [13]. Section III introduces the proposed data compression models and their training procedures using the ultrasonic experimental data. Section IV presents the performance evaluation of the compression algorithms, as well as evaluating the impact of data quantization for further data reduction. Section V concludes the article.

II. Ultrasonic Datasets for Benchmarking

Sections II-A and II-B introduce two ultrasonic datasets for compression analysis. The dataset presented in Section II-A contains ultrasonic backscattered echoes acquired by NDT of a steel block. Section II-B presents ultrasonic medical images available in the public domain, called the OASBUD dataset [13]. The NDT and OASBUD datasets are used to evaluate the compression efficiency and reconstruction signal fidelity.

A. Ultrasonic NDT Steel Block 2-D-Scan Dataset

The ultrasonic NDT dataset contains ultrasonic backscattered echoes from the microstructure of a steel block with a dimension of 100 × 100 × 224 mm^3, using a 5-MHz piezoelectric transducer and sampling frequency of 100 MHz. The acquired ultrasonic backscattered signals can be used for grain size estimation and flaw detection [14], [15]. Two stepper motors in the system allow 1-D or 2-D scans. For the 2-D scan, we acquire 128 × 128 A-Scans representing information on volumetric scattering within the steel block. Fig. 2 shows the plot of a randomly selected ultrasonic A-Scan from the NDT dataset and its power spectral density (PSD). The measured A-Scan represents the microstructure (grains) scattering within the steel block [16], [17]. This is a typical ultrasonic scattering signal consisting of multiple interfering echoes with random amplitudes and arrival times. The PSD of the signal indicates that the spectrum of the signal has a center frequency of 5 MHz. Each A-Scan contains 2048 data points. Fig. 3 shows five B-Scans plotted using the filtered

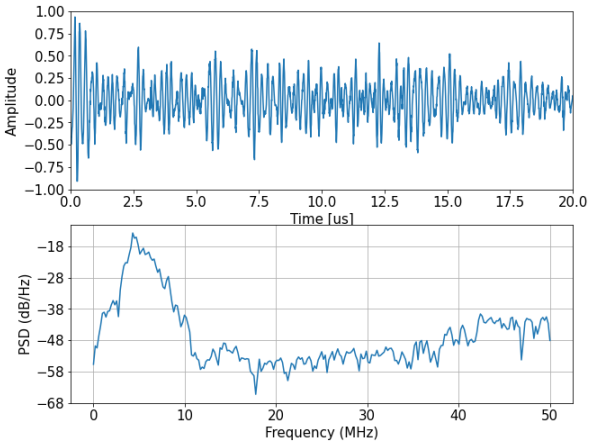


Fig. 2. Ultrasonic NDT backscattered A-Scan from a steel block and the PSD. The dataset is acquired using an immersive ultrasonic transducer centered at 5 MHz with a sampling frequency of 100 MHz.

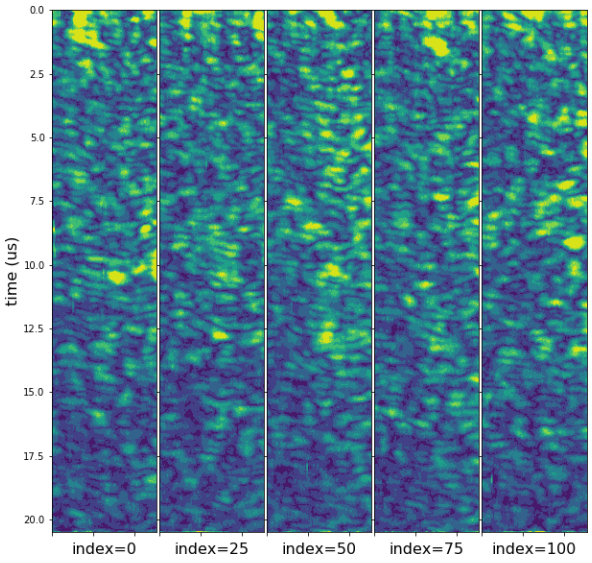


Fig. 3. Filtered 2-D contour plots of the ultrasonic NDT B-Scan envelopes.

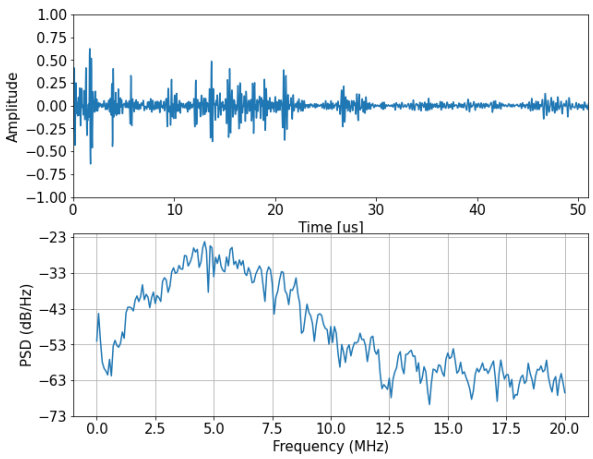


Fig. 4. Ultrasonic medical backscattered A-Scan and the PSD. The dataset is acquired using an ultrasonic linear array transducer centered at 5 MHz with a sampling frequency of 40 MHz.

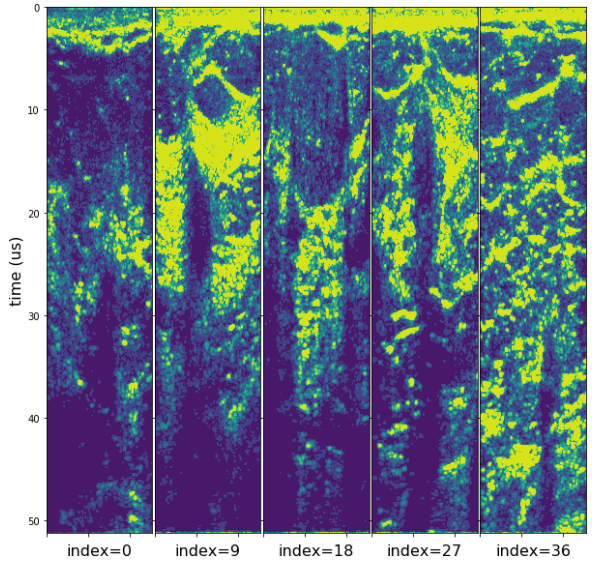


Fig. 5. Filtered 2-D contour plots of the ultrasonic medical B-Scan envelopes.

2-D contour plot function provided in Matplotlib [18]. The B-Scan covers 60.83 mm in the depth of the steel block since the sound speed in steel is 5940 m/s [19]. Before plotting the B-Scans, the amplitude of each A-Scan has been preprocessed by the Hilbert transform to reveal its envelope. Similar results can be obtained in the medical imaging applications for tissue characterization.

B. Ultrasonic OASBUD Medical Image Data Set

The OASBUD dataset represents ultrasonic backscattered echoes acquired with a 40-MHz sampling frequency from breast lesions of patients [13]. We chose 45 ultrasonic B-Scans from the OASBUD dataset to test the proposed compression algorithms; each B-Scan contains 510 A-Scans. Fig. 4 shows a randomly selected A-Scan and its PSD from the OASBUD dataset. Fig. 5 shows five B-Scans of the OASBUD dataset plotted using the same processing procedure, as used in Fig. 3. The collected B-Scans cover 39.4 mm in the depth of the breast tissue if we assume that the speed of sound in tissue is 1540 m/s. Comparing Fig. 5 against Fig. 3 reveals that the OASBUD dataset has a higher degree of texture variation in the B-Scans than the steel block dataset.

III. Training Data Compression Models Using Ultrasonic Experimental Data

This section introduces four unsupervised learning models for data compression algorithms: AE, CAE, WPTCAE, and a WPTCAE/AE hybrid model. These compression models typically consist of two components: an encoder for data compression and a decoder for data reconstruction. These compression algorithms are regarded as unsupervised learning methods since they do not need explicit labels during training. Section III-A introduces AE and CAE algorithms. Section III-B presents the WPTCAE model for ultrasonic data compression. The structure of the WPTCAE model is designed based on the conventional WPT compression

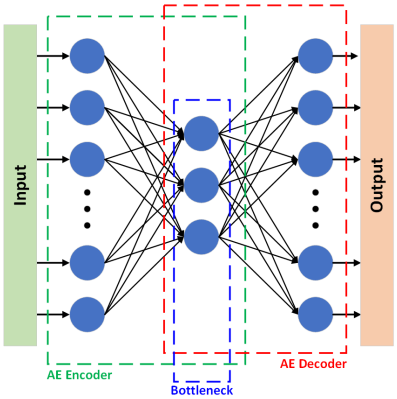


Fig. 6. AE architecture consisting of an encoder and a decoder where the intersection is the bottleneck.

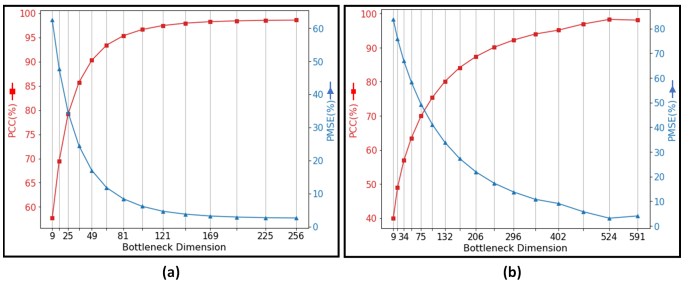


Fig. 7. AE bottleneck dimension versus the PMSE (%) and PCC (%). (a) AE models are trained with the NDT dataset, which consists of 128 × 128 A-Scans. (b) AE models are trained with the OASBUD dataset, which contains 510 × 45 A-Scans.

algorithm. A Python program is implemented to find the optimal WPT compression scheme for both the NDT and the OASBUD datasets. Section III-C introduces a hybrid model that combines the AE and WPTCAE models, a combination referred to as the WPTCAE/AE, for improved compression accuracy.

A. Autoencoder and Convolutional Autoencoder

Fig. 6 shows a typical AE model consisting of an encoder and a decoder. The output of the AE encoder at the bottleneck is the compressed representation of the data (size noted as Q). The AE decoder reconstructs the compressed data into its original (size noted as P). For this study, the AE uses only linear activation functions, and the loss function is the mean square error (mse). The AE model will learn a linear transformation that projects the data onto another space with fewer dimensions, and this is equivalent to a principal component analysis (PCA) model. Training an AE model for compression minimizes the mse between the input and the output using the backpropagation algorithm [20].

One of the major drawbacks of using an AE for compression is the huge number of weights in the AE model (see Fig. 6). In AE compression algorithms, the decoder weights need to be saved for data reconstruction. Assume that we want to build a model with an input dimension of P, one encoder layer with Q neurons (i.e., Q is the size of the bottleneck), and one decoder layer with P output neurons. Then, this model has $2PQ+P+Q$ floating-point (i.e., 4-byte representation) numbers for weights and biases. The model generates a compressed form of the signal with a dimension of $Q \ll P$. If the dataset to be compressed contains K sets of A-Scans, we store only the encoder output and the decoder weights as the compressed data. The original data contain KQ floating-point numbers, whereas the compressed data contain $PY+P+KQ$ floating-point numbers. The compression ratio is, thus, $(KQ/(Q(P+1+K)))$. For practical compression, the compression ratio must be greater than 1. Assume that $P = 2048$ and $Q = 128$ and that we want to achieve a compression ratio of 15. Then, K must be greater than 30 960, which implies that we need a massive amount of ultrasonic data (over 252 MB) for efficient data compression. This is usually the case when ultrasound is used for medical imaging or NDT applications.

A critical parameter of the AE compression model is the bottleneck dimension Q. A properly chosen Q ensures optimal compression accuracy and compression ratio. To determine the optimal bottleneck dimension, Q, we train the AE models with different bottleneck dimensions for the NDT and OASBUD datasets using the Adam optimizer [21] for 50 epochs. The average PMSE and PCC are recorded and plotted against the bottleneck dimensions, as shown in Fig. 7. By inspecting the results in Fig. 7, we set the bottleneck dimension of the AE model at 225 for the ultrasonic NDT dataset (2048 for each A-Scan) and at 512 for the OASBUD ultrasonic medical imaging dataset (2048 for each A-Scan). The results in Fig. 7 show that the AE model can successfully compress both ultrasonic datasets. It also reveals that the AE model is generally not a good option when the dataset to be compressed has a small volume since the compression ratio is lowered because of the overhead resulting from the AE decoder weights needed for signal reconstruction.

An alternative method for reducing the AE decoder weights is the CAE model, as shown in Fig. 8. In general, the CAE model includes one or more convolution layers before the AE encoder and after the AE decoder. After each convolution layer (filters = 1, kernelsize = 100, and stride = 2) in the CAE encoder, the signal is downsampled by a factor of two to reduce the signal dimension. Before each convolution layer in the CAE decoder, the signal is upsampled by the factor of two to restore its original dimension. Each convolution layer (filters = 1, kernelsize = 100, and stride = 2) in the CAE encoder reduces the number of weights by a factor of two compared with an AE model with the same bottleneck dimensions.

B. WPT Optimized by CAEs

Wavelet transformation is a highly practical time–frequency analysis method for representing finite, nonstationary, and nonperiodic signals [22], [23]. Using WPT for data compression has proved to be effective because of the excellent energy compaction and time localization properties of the wavelet kernels [24]. Fig. 9 illustrates one level of the WPT decomposition and reconstruction filter pairs using a wavelet kernel from the Daubechies wavelet family [25]. The low-pass

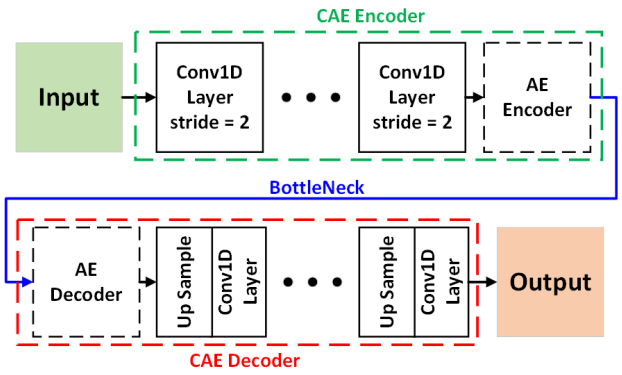


Fig. 8. General block diagram of the CAE.

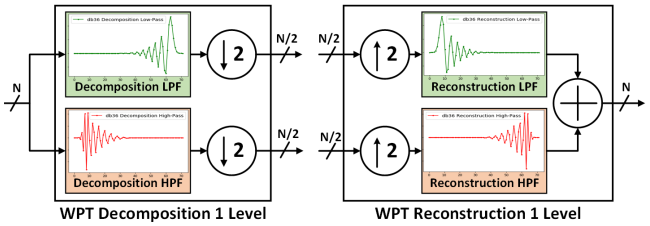


Fig. 9. WPT decomposition and reconstruction filter pairs using a wavelet kernel from the Daubechies wavelet family as an example.

filters (LPFs) and high-pass filters (HPFs) for signal decomposition are known as analysis filters, while the LPFs and HPFs for signal reconstruction are known as synthesis filters. For orthogonal wavelet kernels, the synthesis filters are the time-reversed versions of the analysis filters. The multilevel WPT of a signal is calculated by passing it through a series of decomposition LPFs and HPFs. According to Nyquist's rule, the filtered signals are downsampled by a factor of two because of the filtering half of the frequencies of the signal [25]. After downsampling, the outputs of the LPF (named L) and HPF (named H) are referred to as approximated and detailed components, respectively. The decomposition processing block in WPT is composed of multiple layers of such filter pairs (e.g., the subbands at WPT decomposition Level 2 are named LL, LH, HL, and HH). The reconstruction processing block in the WPT model is the reverse of the decomposition processing block. Before the subbands are passed through the synthesis filters in the WPT reconstruction, the signals are upsampled by inserting zeros at alternate sample points.

When WPT is used for data compression, the input signal is decomposed into 2^N subbands, where N is the number of decomposition layers. We can determine the number of decomposition layers by observing the energy concentration in each subband in multilayer decomposition [5]. The subbands with an energy below a predefined threshold can be eliminated. This action results in lossy data compression with a minimal adverse impact. Consequently, we choose the top M subbands among the 2^N subbands that contain the highest energy level as the compressed data. The compressed data can be further quantized to reduce the number of bits representing each data point; the quantization bits can be determined by the energy associated with that subband. The signal can be reconstructed by applying the inverse WPT (IWPT) to the compressed data.

The success of the WPT compression algorithm highly depends on the selection of the wavelet kernel. In this study, we developed a procedure for finding the optimal wavelet kernel and the optimal WPT decomposition tree structures based on the characteristics of the ultrasound signals. The best energy compaction can be achieved by choosing a wavelet kernel that has the highest similarity with the ultrasonic echoes. In this article, we obtain the optimal wavelet kernels using the PyWavelets library [26]. There are 14 wavelet families, including 106 discrete wavelet filter kernels provided by the library that can be used for WPT decomposition and reconstruction. A Python script is written to iterate through all the discrete wavelet filters to find the optimal kernel for the ultrasonic datasets, as described in Algorithm 1.

Algorithm 1 Optimal Wavelet Kernel for a Given Number of WPT Decomposition Layers, N, and the Number of Most Dominant Decomposition Subbands, M for Reconstruction

1 Randomly select a finite number of A-Scans from the dataset; `// 500 A-Scans is recommended`
2 Select the number of WPT decomposition layers noted as N;
3 Select the number of dominant WPT decomposition subbands for signal reconstruction noted as M;
4 End of initialization step;
`/* Iterate 106 wavelet kernels in PyWavelet Library */`
5 **for** *WPTKernel=Kernel_0:Kernel_105* **do**
6 **for** *DecompositionLevel=1:N* **do**
7 Load WPT lowpass and highpass filters with the current WPTKernel;
8 Take components from the current level, apply WPT lowpass filter and highpass filter to the signals to acquire approximation and detail coefficients;
9 Downsample the output components by a factor of 2 to remove the redundant information;
10 **end for**
11 Compute the energy ratio of each component at target level N;
12 Record the energy distribution at WPT decomposition level N;
13 **end for**
14 Sort the energy distribution of each wavelet kernel;
15 From the sorted energy distribution, find the *Optimal_Kernel* that can preserve the most energy with M decomposition components with highest energy ratio;
16 **return** *Optimal_Kernel*

Algorithm 1 is applied to the ultrasonic NDT and OASBUD datasets described in Section II to estimate the optimal wavelet kernel for each dataset. In Step 11 of Algorithm 1, the energy of the signal in a subband is computed as the sum of the squares of the amplitudes. The optimal wavelet kernel is

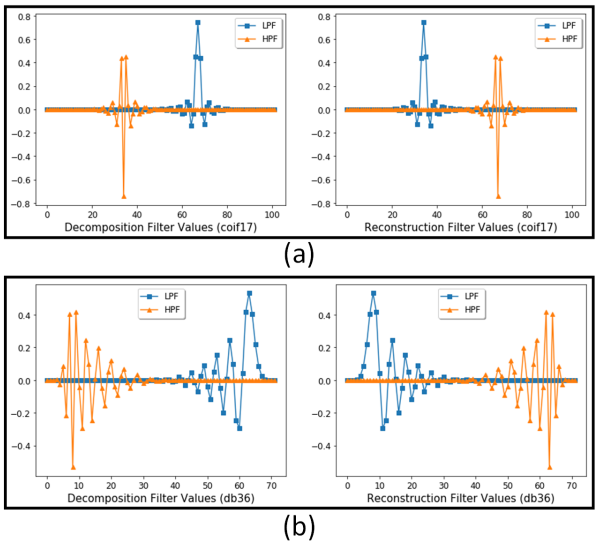


Fig. 10. WPT filter decomposition and reconstruction analysis and synthesis filter kernels. (a) "coif17" from the Coiflet wavelet family. (b) "db36" from the Daubechies wavelet family.

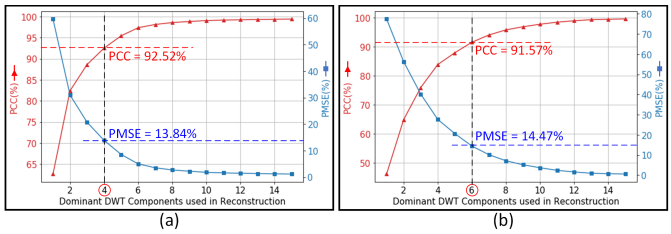


Fig. 11. WPT compression accuracy (PCC and PMSE) versus the number of the most dominant subbands used for signal reconstruction. (a) NDT dataset. (b) OASBUD dataset.

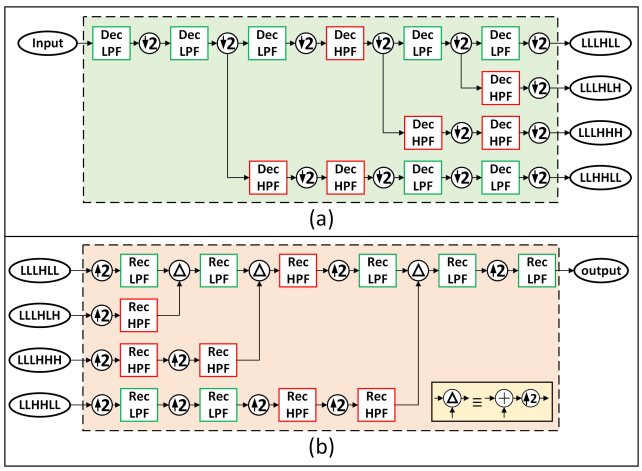


Fig. 12. Optimal WPT compression tree structure for the ultrasonic NDT dataset. (a) WPT decomposition (six levels), with subbands LLLHHH, LLLHLH, LLLHLL, and LLHHLL preserved as the compression output. (b) WPT reconstruction.

chosen from the kernels that can preserve 85%–90% of the total energy with the minimum number of subbands. When $N = 6$ and $M = 4$, the optimal wavelet kernel is "coif17" for the NDT dataset. The WPT decomposition subbands LLLHHH, LLLHLH, LLLHLL, and LLHHLL preserve 85.6% of the overall energy at Level 6 of the WPT decomposition. For the OASBUD ultrasonic medical dataset, "db36" is found to be the optimal wavelet kernel with $N = 5$ and $M = 6$. The WPT decomposition components LLHLL, LHHLL, LLHLH, LHHLH, LLHHH, and LHHHH can preserve 85.24% of the overall energy at WPT decomposition Level 5. Fig. 10 shows both decomposition and reconstruction filter kernels of "coif17" from the Coiflet wavelet family and "db36" from the Daubechies wavelet family.

Fig. 11 shows how the value of M affects the compression accuracy in the WPT compression algorithm for the NDT and OASBUD datasets. In Fig. 11(a), six WPT decomposition levels are applied to the NDT dataset; the x-axis is the number of components preserved in the compression. After WPT reconstruction with the optimal M components, the average PCC [see (2)] and PMSE [see (1)] are computed. From this result, we decided to use four WPT components out of 64 as the compressed output for the NDT dataset. This ensures that the average PCC accuracy is more than 90%. Fig. 11(b) shows a similar result for the OASBUD dataset. The result suggests that six WPT components out of 32 ensure 90% PCC accuracy.

Fig. 12 illustrates the optimal design of the WPT compression and reconstruction processing blocks for the ultrasonic NDT dataset, with "coif17" used as the wavelet kernel. As shown in Fig. 12(a), the subbands LLLHHH, LLLHLH, LLLHLL, and LLHHLL are the compression output. Fig. 12(b) shows the WPT reconstruction procedures for recovering the original data.

Fig. 13 demonstrates the optimal WPT compression and reconstruction processing blocks for the OASBUD dataset, with "db36" used as the wavelet kernel. The subbands LLHLL, LHHLL, LLHLH, LHHLH, LLHHH, and LHHHH at WPT decomposition Level 5 of the OASBUD dataset are preserved as the compression result.

In summary, WPT is effective for ultrasonic data compression. The WPT for compression and reconstruction resembles the CAE structure. Consequently, the architecture of the WPT decomposition and reconstruction processing blocks can be implemented as a CAE model with predefined wavelet kernels. Hence, the filters in WPT compression are replaced by 1-D convolution layers, as shown in Fig. 8. This approach is called WPTCAE and can be trained with the backpropagation algorithm to improve signal reconstruction accuracy. Therefore, we realized the WPTCAE encoder and decoder models for ultrasonic signals based on the WPT decomposition and reconstruction processing blocks shown in Figs. 12 and 13. In the WPTCAE encoder model, the downsampling is achieved by setting the stride to two. In the WPTCAE decoder model, each input of the convolution layer is upsampled by inserting zeros at every other point. The ultrasonic NDT and OASBUD datasets are used as the input training data, and the prediction output of the WPTCAE model is the reconstructed signal. The backpropagation algorithm is used to update the weights and reduce the mse loss between the reconstructed data and the training data. After training, the WPTCAE model is separated into a trained encoder for data compression and a trained decoder for reconstruction. Hence, the trained decoder

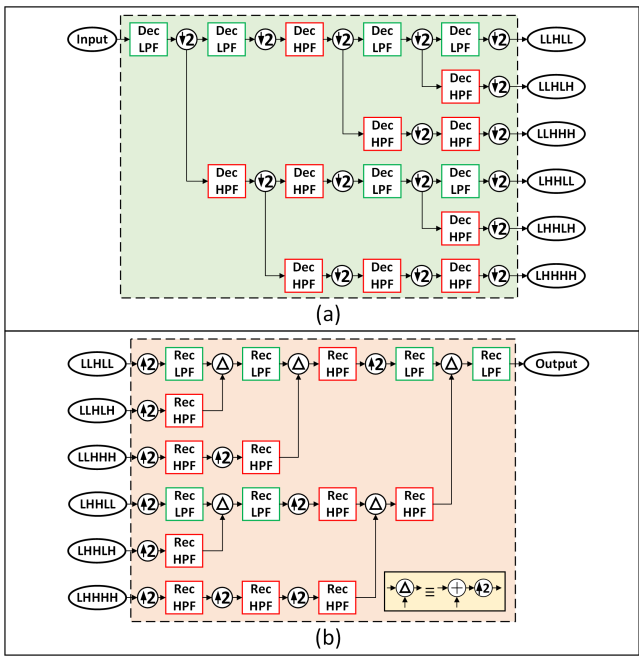


Fig. 13. Optimal WPT compression tree structure for the ultrasonic OASBUD dataset. (a) WPT decomposition (five levels), with subbands LLHLL, LHHLL, LLHLH, LHHLH, LLHHH, and LHHHH preserved as the compression output. (b) WPT reconstruction.

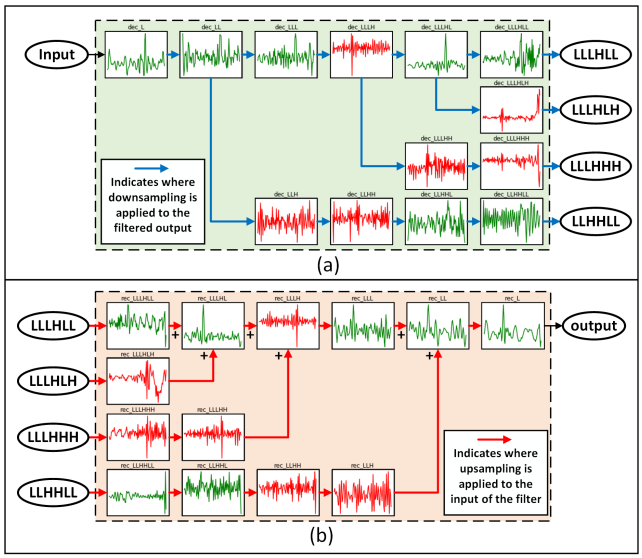


Fig. 14. Trained encoder and decoder in the WPTCAE model for the compression of the ultrasonic NDT dataset. (a) Trained WPTCAE encoder for signal compression. (b) Trained WPTCAE decoder for signal reconstruction.

model must be saved along with the compressed data for reconstruction. Furthermore, it is also desirable to save the trained encoder model for compressing datasets of similar quality. Figs. 14 and 15 exhibit the trained WPTCAE models applied to Figs. 12 and 13, respectively. The initial values of the filter kernels can be found in Fig. 10 for both datasets.

C. WPTCAE/AE Hybrid Model

The WPTCAE model is developed based on the WPT compression algorithm, an optimized filter combination for

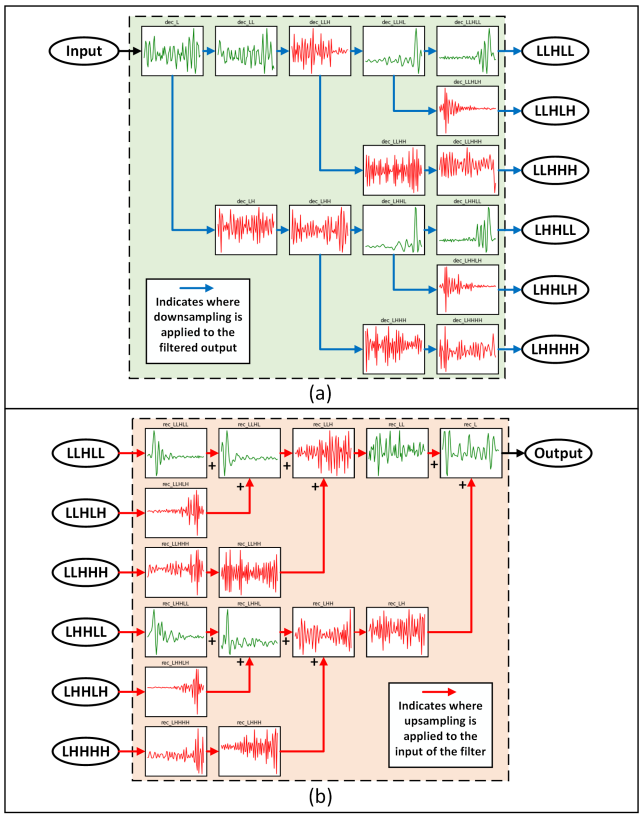


Fig. 15. Trained encoder and decoder in the WPTCAE model for the compression of the ultrasonic OASBUD medical image dataset. (a) Trained WPTCAE encoder for signal compression. (b) Trained WPTCAE decoder for signal reconstruction.

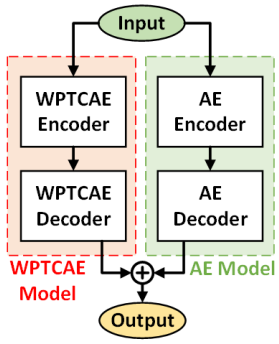


Fig. 16. Block diagram of the WPTCAE/AE hybrid model.

preserving the information in specific frequency bands. A certain amount of information is unavoidably lost if we do not use all the decomposed subbands for signal reconstruction. The amount of lost information can be minimized by including a complimentary processing block that recovers losses. If the reconstructed signal fidelity is critical, one can complement the WPTCAE model with an AE model, a combination referred to as the WPTCAE/AE hybrid model, to improve signal reconstruction accuracy. This hybrid model is shown in Fig. 16, in which the input signal is broadcast into both the WPTCAE and the AE model, and the output is the sum of the AE and WPTCAE predictions. For efficiency, the bottleneck dimension of the AE is designed to be relatively low to reduce the compression overhead caused by the AE model.

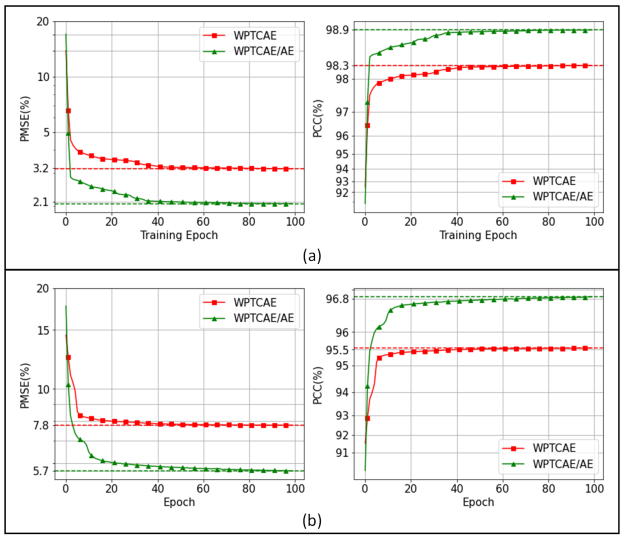


Fig. 17. Comparison of the training history (PCC and PMSE) of the WPTCAE and WPTCAE/AE compression models. (a) NDT dataset. (b) OASBUD dataset.

We use the ultrasonic datasets to train the WPTCAE models and the WPTCAE/AE hybrid models for comparison. In the hybrid models for the NDT and OASBUD datasets, the bottleneck dimension of the AE models is 16. For the ultrasonic NDT dataset, Fig. 17(a) shows the training history (PCC and PMSE) of 100 epochs of the WPTCAE model compared with the WPTCAE/AE hybrid model. After training, the final PMSE and PCC values show that the trained hybrid model has a higher compression accuracy than the WPTCAE model. During the training, the convergence rate of the WPTCAE/AE hybrid model is superior to the convergence rate of the WPTCAE model. With the current data, Fig. 17(a) and (b) indicates that the optimal solution can be reached within 50 epochs. Also, the reconstruction accuracy of the signal is increased by 0.5% for the ultrasonic NDT dataset and 1.3% for the OASBUD medical image dataset.

IV. Results and Analysis

This section presents the performance evaluation of the proposed compression algorithms using the ultrasonic experimental data. The compression performance is evaluated by the compression accuracy measured by the PCC and PMSE and the compression efficiency measured by the compression ratio. Section IV-A presents the compression performance of the proposed algorithms, including the AE, CAE, WPT, WPTCAE, and WPTCAE/AE models. By applying the optimal bit-width quantization to the compression output, one can further enhance the compression efficiency while maintaining a comparable signal reconstruction accuracy. Section IV-B presents the experimental results when quantization is incorporated into the WPTCAE compression models.

A. Compression Results and Performance Analysis

In the analysis of the compression performance, the input data, compressed data, reconstructed data, and model weights are all considered to be 32-bit floating-point numbers. The search for the optimal data compression models is achieved through 150 epochs of training. Table I shows the compression performance of the algorithms applied to the ultrasonic NDT dataset. According to Fig. 7(a), the optimal AE bottleneck dimension is 225 for the NDT dataset. The CAE compression model has one convolution layer with filters $=1$, kernelsize $=100$, and stride $=2$ in the encoder and one convolution layer with filters $=1$, kernelsize $=100$, and upsampling by a factor of 2 in the decoder. The bottleneck dimension of the CAE is also 225. The WPTCAE model structure applied to the ultrasonic NDT dataset is shown in Fig. 12. The weights of the convolution layers are initialized with the "coif17" wavelet kernel. The bottleneck dimension of the AE in the WPTCAE/AE hybrid model is 16, and this arrangement requires an additional 67 600 trainable weights compared with the WPTCAE model. When the hyperparameters are optimally estimated, both the AE and CAE models exhibit outstanding compression accuracies (PCC) of 98.63% and 99.04%, respectively. However, the AE and CAE compression models contain a huge number of weights, resulting in a lower compression ratio.

The conventional WPT compression algorithm contains a minimal number of predefined weights and, consequently, has the lowest compression accuracy. However, the WPTCAE model, after training, has significantly better compression accuracy. When signal fidelity is critical, the WPTCAE/AE model is a desirable choice for the best compression accuracy. The additional overhead in the WPTCAE/AE hybrid model due to the AE size is insignificant for massive data compression.

Table II shows the compression performance of the proposed algorithms for the OASBUD ultrasonic medical image dataset. The bottleneck dimension of the AE model is 512, according to the result in Fig. 7(b). The CAE has one convolution layer with filters $=1$, kernelsize $=100$, and stride $=2$ in the encoder and one convolution layer with filters $=1$, kernelsize $=100$, and upsampling by a factor of 2 in the decoder. The bottleneck dimension of the CAE is also 512. The WPTCAE model structure applied to the ultrasonic OASBUD medical image dataset is shown in Fig. 13. The bottleneck dimension of AE in the WPTCAE/AE hybrid model is 16, and this arrangement requires an additional 67 600 trainable weights compared with the WPTCAE model. The result in Table II confirms the superior performance of the WPTCAE model in terms of compression accuracy and compression ratio.

Both the NDT and OASBUD datasets occupy a similar frequency range from 0 to 10 MHz (as shown in Figs. 2 and 4). The NDT dataset is sampled at 100 MHz, while the OASBUD dataset is sampled at 40 MHz. This means that the NDT dataset is more oversampled and more redundant compared with the OASBUD dataset. Consequently, one expects a higher compression ratio for the ultrasonic NDT dataset than for the OASBUD dataset. This can be clearly observed by inspecting the compression ratios in Tables I and II; the NDT dataset has a compression ratio that is at least two times better than the compression ratio of the OASBUD dataset.

TABLE I

NDT DATASET COMPRESSION PERFORMANCE COMPARISON

	AE-225	CAE-225	WPT	WPTCAE	WPTCAE/AE
Number of Weights	923,873	462,251	200	2,652	70,252
Bottleneck Dimension	225	225	128	128	144
Overall Data	4,610,273	4,148,651	2,097,352	2,099,804	2,429,548
Compression Ratio	7.28	8.09	16	15.97	13.81
PCC(%)	98.63	99.04	92.52	98.36	98.9
PMSE(%)	2.57	1.77	13.85	3.06	2.04

TABLE II

OASBUD DATASET COMPRESSION PERFORMANCE COMPARISON

	AE-512	CAE-512	WPT	WPTCAE	WPTCAE/AE
Number of Weights	2,099,712	1,050,314	140	2,160	69,760
Bottleneck Dimension	512	512	384	384	400
Overall Data	13,850,112	12,800,714	8,812,940	8,814,960	9,249,760
Compression Ratio	3.40	3.67	5.33	5.33	5.08
PCC(%)	97.21	98.65	91.57	95.6	96.9
PMSE(%)	5.54	2.13	14.47	7.70	5.56

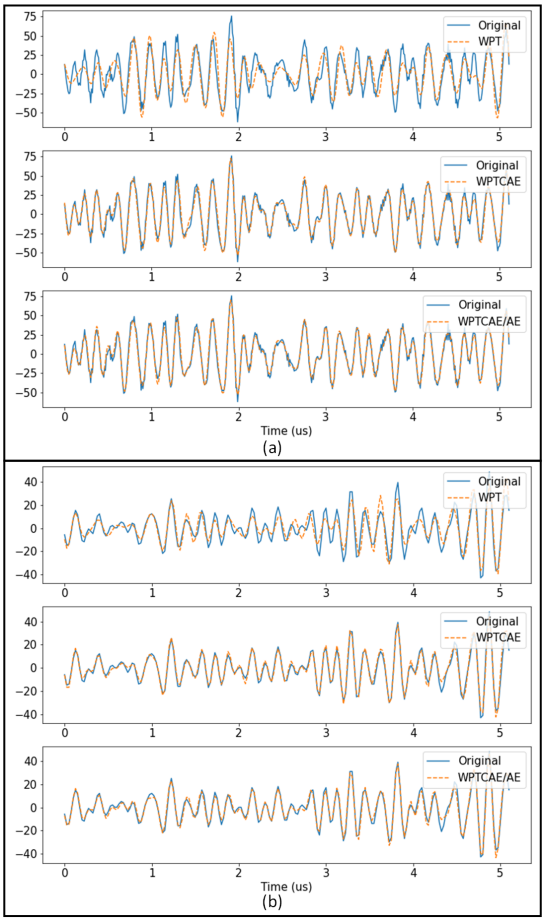


Fig. 18. Comparison of the original A-Scans and the reconstructed signals using the WPT, WPTCAE, and WPTCAE/AE compression models. (a) NDT dataset. (b) OASBUD dataset.

Fig. 18 shows the time-domain comparison of randomly selected A-Scans from the ultrasonic NDT and OASBUD datasets and their reconstructions. This figure shows the time-domain performance comparison between the original and reconstructed signals using the WPT, WPTCAE, and WPTCAE/AE compression models. The WPTCAE

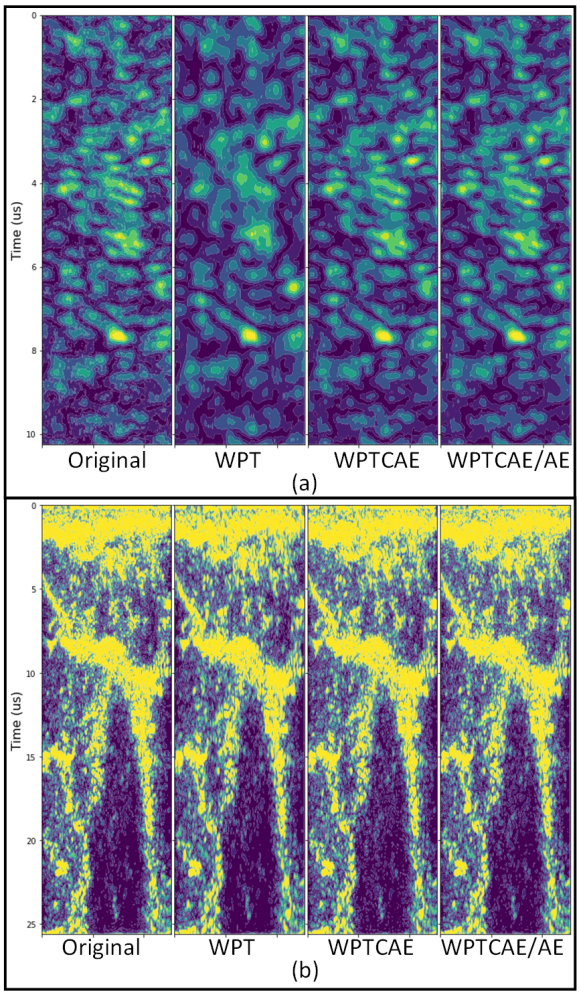


Fig. 19. Comparison of the original B-Scans and the reconstructed B-Scans using the WPT, WPTCAE, and WPTCAE/AE compression models. (a) NDT dataset. (b) OASBUD dataset.

compression model displays an extraordinary compression accuracy compared with the conventional WPT algorithm. Furthermore, this figure also shows that the WPTCAE/AE model offers the best reconstructed signal accuracy for both the NDT and OASBUD datasets. It is important to point out that wavelet-based compression algorithms are designed using selected low-frequency subbands covering the frequency range of ultrasonic used in the measuring system [4], [5]. Subbands associated with higher frequencies are outside the frequency range of the measuring system and represent instrumentation noise rather than ultrasonic scatter echoes. Therefore, compression algorithms not only compress the data but also improve the signal-to-noise ratio. This phenomenon is evident in experimental results presented in Figs. 18 and 19.

Fig. 19 shows the comparison of the original B-Scans and the reconstructed B-Scans processed by the WPT, WPTCAE, and WPTCAE/AE compression algorithms. In this figure, one can observe an apparent difference between the original B-Scans and the reconstructed B-Scans when the WPT compression is used. Some details are missing from the WPT-reconstructed B-Scans because of the algorithm's

TABLE III

WPTCAE COMPRESSION PERFORMANCE ON THE NDT DATASET WITH BIT PRECISION AND QUANTIZATION

	WPTCAE (int8)	WPTCAE with Quantization
Original Data Size	32 MB	32 MB
Compressed Data Size	2 MB	1.375 MB
WPTCAE Model Size	10.37 kB	10.37 kB
Compression Ratio	15.92	23.3
PCC(%)	98.22	96.58
PMSE(%)	3.34	6.49

TABLE IV

WPTCAE COMPRESSION PERFORMANCE ON THE OASBUD DATASET WITH BIT PRECISION AND QUANTIZATION

	WPTCAE (int8)	WPTCAE with Quantization
Original Data Size	44.82 MB	44.82 MB
Compressed Data Size	8.4 MB	5.25 MB
WPTCAE Model Size	8.44 kB	8.44 kB
Compression Ratio	5.33	8.54
PCC(%)	94.67	92.54
PMSE(%)	10.65	14.29

low compression accuracy since a limited number of subbands are used in the image reconstruction. As expected, the best compression results are achieved by the WPTCAE and WPTCAE/AE models.

B. Quantization and Bit Precision

Another important aspect of the compression algorithms is the data bit precision. This section discusses the importance of quantization and bit precision in the WPTCAE models. The ultrasonic NDT and OASBUD 32-bit floating-point datasets are scaled and quantized into 8-bit precision numbers ranging from -128 to 127. The weights of the WPTCAE models remain as 32-bit floating-point numbers for higher processing precision for decomposition and reconstruction. The subbands of the compressed output are quantized into 8 bits or fewer in order to increase the compression ratio.

For the WPT compression algorithms, the quantization can be directly applied to the decomposed subbands according to the energy distribution. For the NDT dataset, the four subbands LLLHHH, LLLHLH, LLLHLL, and LLHHLL are preserved as the compressed output, in which the energies associated with these four subbands are 34.73%, 25.65%, 11.67%, and 9.28%, respectively. These four subbands are quantized into 7, 6, 4, and 3 bits using the energy distribution of the subbands as the guiding factor. As a result of the quantization, the average PCC and PMSE values (when the original data, compressed data, and reconstructed data are 8-bit numbers) are slightly sacrificed from 26.96% and 92.52% to 28.67% and 91.55%, respectively. However, with this quantization, the compression ratio is improved by 37.75%. As shown in Table III, a similar performance improvement can be achieved by adding quantization layers in the trained WPTCAE compression model. After training, the energy distributions of the WPTCAE subbands for the NDT dataset are 17.89%, 10.78%, 33.82%, and 37.51%. Accordingly, these four subbands are quantized into 5, 4, 6, and 7 bits.

Table IV shows the performance comparison when quantization is added to the WPTCAE model for the OASBUD dataset. After training, the energy distributions of the six WPTCAE subbands are 11.19%, 9.76%, 34.18%, 9.57%, 10.26%, and 25.04%. Accordingly, these six subbands are quantized into 4, 4, 7, 4, 4, and 7 bits. The results in Table III and Table IV suggest that adding quantization to the trained WPTCAE model can increase the compression ratio significantly, while the reduction in compression accuracy is negligible. Fig. 20 shows the time-domain comparison of the original A-Scans and the reconstructed A-Scans when quantization is applied

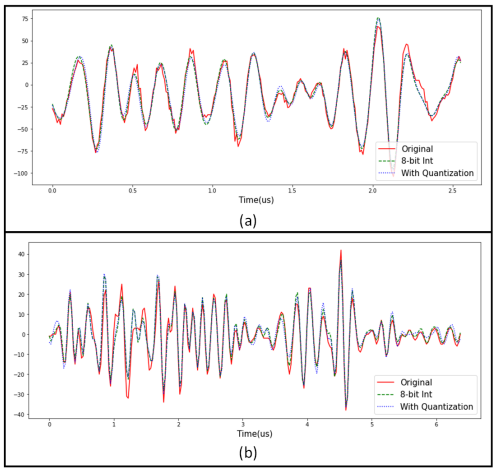


Fig. 20. Comparison between the original A-Scans and the reconstructed signals of the WPTCAE compression models when quantization is introduced into the algorithm. (a) NDT dataset, with the subbands quantized into 5, 4, 6, and 7 bits. (b) OASBUD dataset, with the subbands quantized into 4, 4, 7, 4, 4, and 7 bits.

to the WPTCAE models. In this figure, the original A-Scans are plotted in red solid lines, the A-Scans reconstructed from the 8-bit subbands are in green dashed lines, and the A-Scans reconstructed from the quantized subbands are in blue dotted lines.

V. CONCLUSION

This article presents ultrasonic data compression algorithms, including AE, CAE, DWTCAE, and DWTCAE/AE learning models. These algorithms are data-adaptive and offer optimal solutions for compressing ultrasonic NDT and ultrasonic medical imaging datasets. The design, tradeoff, and performance of these algorithms are explored. The AE and CAE models for ultrasonic signal compression achieve high compression accuracy when the models' hyperparameters are optimally designed. The practical use of the AE and CAE models for compression is limited because both models generate a huge number of weights that will reduce the compression ratio. The WPTCAE compression models are designed by adopting and optimizing the WPT compression architecture designed for compressing ultrasonic signals. The WPTCAE model is robust and offers extraordinary performance in terms of compression accuracy and compression ratio. The WPTCAE/AE hybrid model is an improved version of the WPTCAE model complemented with a compact AE for higher compression accuracy. Finally, the impact of applying adaptive quantization to the WPTCAE model is explored for an even higher compression ratio.

REFERENCES

[1] D. Park and M. H. Lee, "Image compression based on best wavelet packet bases," *IEEE Trans. Consum. Electron.*, vol. 40, no. 3, pp. 527–537, Aug. 1994.

[2] X. Zha, R. Fu, Z. Dai, and B. Liu, "Noise reduction in interferograms using the wavelet packet transform and Wiener filtering," *IEEE Geosci. Remote Sens. Lett.*, vol. 5, no. 3, pp. 404–408, Jul. 2008.

[3] G. Bhatnagar, B. Raman, and Q. M. J. Wu, "Robust watermarking using fractional wavelet packet transform," *IET Image Process.*, vol. 6, no. 4, pp. 386–397, Jun. 2012.

[4] P. Govindan and J. Saniie, "Processing algorithms for three-dimensional data compression of ultrasonic radio frequency signals," *IET Signal Process.*, vol. 9, no. 3, pp. 267–276, May 2015.

[5] E. Oruklu, N. Jayakumar, and J. Saniie, "Ultrasonic signal compression using wavelet packet decomposition and adaptive thresholding," in *Proc. IEEE Ultrason. Symp.*, Nov. 2008, pp. 171–175.

[6] F. N. Khan and A. P. T. Lau, "Robust and efficient data transmission over noisy communication channels using stacked and denoising autoencoders," *China Commun.*, vol. 16, no. 8, pp. 72–82, Aug. 2019.

[7] Y. H. Lai, F. Chen, S.-S. Wang, X. Lu, Y. Tsao, and C.-H. Lee, "A deep denoising autoencoder approach to improving the intelligibility of vocoded speech in cochlear implant simulation," *IEEE Trans. Biomed. Eng.*, vol. 64, no. 7, pp. 1568–1578, Sep. 2017.

[8] D. Del Testa and M. Rossi, "Lightweight lossy compression of biometric patterns via denoising autoencoders," *IEEE Signal Process. Lett.*, vol. 22, no. 12, pp. 2304–2308, Dec. 2015.

[9] Z. Cheng, H. Sun, M. Takeuchi, and J. Katto, "Energy compaction-based image compression using convolutional autoencoder," *IEEE Trans. Multimedia*, vol. 22, no. 4, pp. 860–873, Apr. 2020.

[10] Z. Cheng, H. Sun, M. Takeuchi, and J. Katto, "Deep convolutional autoencoder-based lossy image compression," in *Proc. Picture Coding Symp. (PCS)*, Jun. 2018, pp. 253–257.

[11] M. Abadi *et al.* (2015). *TensorFlow: Large-Scale Machine Learning on Heterogeneous Systems*. [Online]. Available: https://www.tensorflow.org/

[12] P. Govindan, B. Wang, P. Ravi, and J. Saniie, "Hardware and software architectures for computationally efficient three-dimensional ultrasonic data compression," *IET Circuits Devices Syst.*, vol. 10, no. 1, pp. 54–61, Jan. 2016.

[13] H. Piotrzkowska-Wróblewska, K. Dobruch-Sobczak, M. Byra, and A. Nowicki, "Open access database of raw ultrasonic signals acquired from malignant and benign breast lesions," *Med. Phys.*, vol. 44, no. 11, pp. 6105–6109, Nov. 2017, doi: 10.1002/mp.12538.

[14] B. Wang and J. Saniie, "A high performance ultrasonic system for flaw detection," in *Proc. IEEE Int. Ultrason. Symp. (IUS)*, Oct. 2019, pp. 840–843.

[15] B. Wang and J. Saniie, "Multilayer perceptron neural networks for grain size estimation and classification," in *Proc. IEEE Int. Ultrason. Symp. (IUS)*, Oct. 2019, pp. 1643–1646.

[16] J. Saniie, T. Wang, and N. M. Bilgutay, "Analysis of homomorphic processing for ultrasonic grain signal characterization," *IEEE Trans. Ultrason., Ferroelectr., Freq. Control*, vol. 36, no. 3, pp. 365–375, May 1989.

[17] J. Saniie and N. M. Bilgutay, "Quantitative grain size evaluation using ultrasonic backscattered echoes," *J. Acoust. Soc. Amer.*, vol. 80, no. 6, pp. 1816–1824, Dec. 1986, doi: 10.1121/1.394296.

[18] J. D. Hunter, "Matplotlib: A 2D graphics environment," *Comput. Sci. Eng.*, vol. 9, no. 3, pp. 90–95, May/Jun. 2007.

[19] The Engineering ToolBox. (2004). *Velocity of Sound in Solids and Metals*. Accessed: Mar. 9, 2021. [Online]. Available: https://www.engineeringtoolbox.com

[20] H. J. Kelley, "Gradient theory of optimal flight paths," *ARS J.*, vol. 30, no. 10, pp. 947–954, 1960.

[21] D. P. Kingma and J. Ba, "Adam: A method for stochastic optimization," 2014, *arXiv:1412.6980*. [Online]. Available: https://arxiv.org/abs/1412.6980

[22] I. Daubechies, "The wavelet transform, time-frequency localization and signal analysis," *IEEE Trans. Inf. Theory*, vol. 36, no. 5, pp. 961–1005, Sep. 1990.

[23] O. Rioul and M. Vetterli, "Wavelets and signal processing," *IEEE Signal Process. Mag.*, vol. 8, no. 4, pp. 14–38, Oct. 1991.

[24] P. Kaur, "Comparative analysis between DWT and WPD techniques of speech compression," *IOSR J. Eng.*, vol. 2, no. 8, pp. 120–128, Aug. 2012.

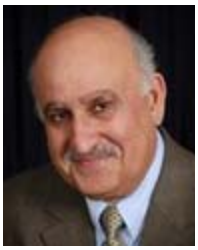
[25] C. Heil, "Ten lectures on wavelets," *SIAM Rev.*, vol. 35, no. 4, pp. 666–669, 1993. [Online]. Available:http://www.jstor.org/stable/2132406

[26] G. R. Lee, R. Gommers, F. Waselewski, and K. Wohlfahrt, "Pywavelets: A Python package for wavelet analysis," *J. Open Source Softw.*, vol. 4, no. 36, p. 1237, 2019, doi: 10.21105/joss.01237.

Boyang Wang (Member, IEEE) received the B.S. degree in information engineering from Beijing Institute of Technology, Beijing, China, in 2013, and the M.S. degree in electrical engineering from Illinois Institute of Technology (IIT), Chicago, IL, USA, in 2015, where he is currently pursuing the Ph.D. degree with the Electrical and Computer Engineering Department.

He is currently a Research Assistant with the Electrical and Computer Engineering Department, IIT. He was a Graduate Student Research Assistant with the Nuclear Science and Engineering Division, Argonne National laboratory, from 2017 to 2018. His research was supported in part by the Walter and Harriet Filter Endowment from 2016 to 2020. His research interests include ultrasonic signal processing, software-defined radio, automation, the Internet of Things, embedded system development, hardware and software codesign, system-on-chip design, artificial intelligence, and machine learning.

Dr. Wang won the Student Paper Competition from the 2019 IEEE International Ultrasonics Symposium. He received the 2020 Sigma Xi/IIT Student Award for Excellence in University Research.



Jafar Saniie (Life Fellow, IEEE) received the B.S. degree (Hons.) in electrical engineering from the University of Maryland, College Park, MD, USA, in 1974, the M.S. degree in biomedical engineering from Case Western Reserve University, Cleveland, OH, USA, in 1977, and the Ph.D. degree in electrical engineering from Purdue University, West Lafayette, IN, USA, in 1981.

In 1981, he joined the Department of Applied Physics, University of Helsinki, Helsinki, Finland, to conduct research in photothermal and photoacoustic imaging. Since 1983, he has been with the Department of Electrical and Computer Engineering, Illinois Institute of Technology, Chicago, IL, USA, where he is the Department Chair, the Filmer Endowed Chair Professor, and the Director of the Embedded Computing and Signal Processing (ECASP) Research Laboratory. He has over 370 publications and has supervised 35 Ph.D. dissertations and 22 M.S. theses to completion. His research interests and activities are in ultrasonic signal and image processing, ultrasonic software-defined communications, artificial intelligence and machine learning, statistical pattern recognition, estimation and detection, data compression, time–frequency analysis, embedded digital systems, system-on-chip hardware/software codesign, the Internet of Things and cybersecurity, computer vision, deep learning, and ultrasonic nondestructive testing and imaging.

Dr. Saniie is an IEEE Life Fellow for contributions to ultrasonic signal processing for detection, estimation, and imaging. He has been a Technical Program Committee Member of the IEEE Ultrasonics Symposium since 1987. He was the Chair of the Sensors, Nondestructive Evaluation (NDE), and Industrial Applications Group from 2004 to 2013 and the Publication Chair of the IEEE Ultrasonics Symposium in 2013 and 2021. He was the General Chair of the 2014 IEEE Ultrasonics Symposium in Chicago. He served as the Ultrasonics Vice President for the IEEE Ultrasonics, Ferroelectrics, and Frequency Control (UFFC) Society from 2014 to 2017. He has been the IEEE UFFC Awards Chair since 2018. He received the 2007 University (Illinois Institute of Technology) Excellence in Teaching Award. He was a Lead Guest Editor for the Special Issue on Ultrasonics and Ferroelectrics of the IEEE TRANSACTIONS ON ULTRASONICS, FERROELECTRICS, AND FREQUENCY CONTROL in August 2014, the Special Issue on Novel Embedded Systems for Ultrasonic Imaging and Signal Processing of the IEEE TRANSACTIONS ON ULTRASONICS, FERROELECTRICS, AND FREQUENCY CONTROL in July 2012, and the Special Issue on Advances in Acoustic Sensing, Imaging, and Signal Processing of the *Journal of Advances in Acoustics and Vibration* in 2013. He has been an Associate Editor of the IEEE TRANSACTIONS ON ULTRASONICS, FERROELECTRICS, AND FREQUENCY CONTROL since 1994.