Reverberation Noise Suppression in the Aperture Domain Using 3D Fully Convolutional Neural Networks

Leandra L. Brickson*, Dongwoon Hyun†, and Jeremy J. Dahl†
Departments of *Electrical Engineering and †Radiology, Stanford University, Stanford, CA 94305
Email: llbricks@stanford.edu

***Abstract*—Reverberation clutter is image noise resulting from multiple reflections between tissue layers and leads to image degradation. We propose a 3D fully convolutional neural network (FCNN) to selectively remove reverberation noise from ultrasound channel data. A training set was generated using Field II Pro to simulate full synthetic aperture channel data. Reverberation noise was approximated by adding bandpass filtered noise to the noise-free simulated channel data. This channel data, with and without noise, was used to train a 3D FCNN with a custom architecture to remove the reverberation and thermal noise. The network was evaluated on a simulated validation set and on an ATS 549 phantom using a L12-3v transducer connected to a Verasonics Vantage 256 system with reverberation noise. The resulting network shows significant reverberation noise reduction. Across the validation set, the normalized RMS loss between the noisy and noise free images decreased from 77% to 39% after being decluttered using 3D FCNN.**

I. Introduction

Reverberation has been shown to be a major source of ultrasound image degradation [1]. Reverberation clutter is caused by multiple reflections between proximal tissue layers [1], confounding the received reflections from distal regions of the tissue. This is especially a problem in high body-mass-index patients, in whom there are many interfaces between fat and connective tissue that are prone to multiple reflections [2]. The resulting image degradation manifests as a loss of contrast in hypoechoic or anechoic regions such as cysts and blood vessels [3]. In addition to degrading B-mode image quality, reverberation clutter diminishes the performance of techniques that require aperture domain signals, such as phase aberration correction, adaptive and alternative beamforming techniques, and sound speed estimation.

Many promising reverberation clutter removal techniques have been proposed, including coupling high frequency pulses onto lower frequency carrier waves [4], filtering the received aperture domain signals via multi-phase apodization with cross-correlation [5] and spatial prediction filtering [6]. Alternatively, a frequency-domain approach has also been proposed using model-based reconstruction [**?**] as well as deep neural networks [7]. In this work, we propose a new technique that uses a 3D fully convolutional neural network to selectively remove reverberation noise from ultrasound channel data in the aperture domain.

II. Methods

A. Simulated Dataset

To train a reverberation noise reduction network, noisy and noise-free channel signals are required for training. To obtain this paired data, a series of images were simulated using the Field II Pro simulation package [8], [9], [10]. Image content was taken from the publicly available ImageNet [11] and Places2 [12] datasets. Although nonsensical for ultrasound imaging, the images provided a wide range of shapes and contrasts for training. Examples of these images are shown in the left column of Figure 1. Images were converted to grayscale and the brightness value of each pixel was assigned as the reflectivity of the local scatterers. Speckle was simulated with random uniformly distributed scatterers in a $10\,\text{mm}\times10\,\text{mm}\times3\,\text{mm}$ phantom centered at the elevation focus of 2 cm. The scatterer density was chosen to be 60 scatterers per resolution cell, and the scattering amplitudes were normally distributed and weighted according to the image reflectivity. A full synthetic aperture dataset was simulated using single-element transmits at 8 MHz with 60% bandwidth to mimic ultrasound channel data from 128 elements of a Verasonics L12-3v linear array transducer. This simulation was run on 1,024 unique images. Examples of B-mode images from of these simulation results are shown in the right column of Figure 1.

Reverberation noise was approximated by adding bandpass filtered noise, centered about the transmit frequency of the array, ranging from -20 to 10 dB of the root-mean-square (RMS) value of the noise-free simulated channel data. Random thermal noise ranging from -20 to 10 dB was also added. We then converted the RF channel data to complex in-phase and quadrature data. Finally, images were resampled as patches in the lateral and axial dimensions, to give 9,000 distinct image-data samples with independent realizations of noise. The image patches had dimensions $64\times64\times128$, where the first two dimensions spanned the lateral and axial image space, and the last dimension spanned the transducer elements. This paired channel data (i.e., with and without noise) composed the training set for the network. A separate validation set was set aside from this dataset of simulations for evaluation.

Fig. 1. Examples of Field II Pro simulation output. The left column is the input echogenicity map, the right column is the B-mode image of the simulated full synthetic aperture data.

B. Neural Network Architecture

A 3D fully convolutional network architecture was chosen to remove reverberation noise. In this network, a series of 3D convolution kernels followed by the non-linear ReLU [13] function were applied to the input to supply an output of the same shape. The fully convolutional architecture implicitly assumes a spatially-localized model of reverberation clutter, enforced by the use of *local* comparisons of lateral, axial, and element data to assess reverberation noise, as opposed to the global approach taken by fully connected networks. A diagram of this network is shown in Figure 2.

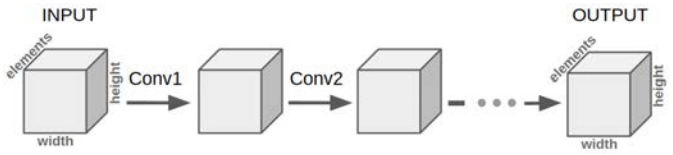

Fig. 2. Illustration of the 3D fully convolutional neural network architecture used for reverberation noise suppression. At each layer, a 3D convolution kernel is applied across the image height, width and transducer element dimensions. Various network depths and hidden layer sizes were tested.

This architecture choice has a few advantages. First, choosing a fully convolutional network instead of having intermediate fully connected layers allows any image size as an input, making the network more versatile to use after training. Additionally, training on only convolution kernels creates significantly fewer parameters for the network to train, which decreases training time and network inference after training, and allows training on smaller datasets.

To train on complex data in TensorFlow, the real and imaginary components were treated as two separate input channels, analogous to the color channels in traditional image analysis networks.

TABLE I
FINAL NETWORK ARCHITECTURE

Layer	Kernel Size	Num Filters
Conv 1 - 2	(3, 3, 11)	5
Conv 3 - 4	(3, 3, 7)	5
Conv 5 - 6	(3, 3, 3)	5
Conv 7	(3, 3, 3)	2

C. Optimization and Evaluation

This network was trained in Python using TensorFlow[14] on the Adam optimizer[15]. For the optimization loss function, either the ℓ_1 or ℓ_2 norm were used as metrics evaluating reconstruction error of the network output:

$$\mathcal{L}_{\ell_1}(y,\hat{y}) = \frac{1}{P}\sum_{p=1}^{P}|y_p - \hat{y}_p| \tag{1}$$

$$\mathcal{L}_{\ell_2}(y,\hat{y}) = \sqrt{\frac{1}{P}\sum_{p=1}^{P}(y_p - \hat{y}_p)^2} \tag{2}$$

Where y_p and $\hat{y}_p$ denote the p-th element in the output. Additionally, the ℓ_2 norm of the network weights was added to the cost function as a regularizer to enforce smaller, more stable convolution kernels.

During training, hyperparameter tuning was performed to determine network depth, hidden layer size, convolutional kernel size, learning rate, batch size, and the ℓ_2 regularization of filter weights. For evaluation of these parameters, the normalized root-mean-square (RMS) loss of the simulated validation set was used at a metric:

$$\text{RMS}_{\text{norm}}(y,\hat{y}) = \mathcal{L}_{\ell_2}(\frac{y_p}{\|y\|}, \frac{\hat{y}_p}{\|\hat{y}\|})| \tag{3}$$

Finally, the network was tested on an ATS 549 phantom using a L12-3v transducer connected to a Verasonics Vantage 256 system. A 0.5 cm thick layer of compressed steel wool was placed between the transducer and phantom to generate reverberation clutter.

III. RESULTS AND DISCUSSION

The final network architecture after tuning is detailed in Table I, where the 'kernel size' is the convolutional layer kernel size and 'num filters' is the hidden layer size of each layer. Using the RMS_{norm} loss as a performance metric, the optimal network length was found to be 7 layers, and the optimal training loss function was found to be the ℓ_1 norm loss. Interestingly, we observed that performance was improved when the 3D convolution kernels in the earlier layers was longer in the dimension of the element signals. Of the tested sizes, the convolutional kernel size found to give best performance was (3,3,11) for the first two layers, (3,3,7) for the next two, and (3,3,3) for the remaining layer. This

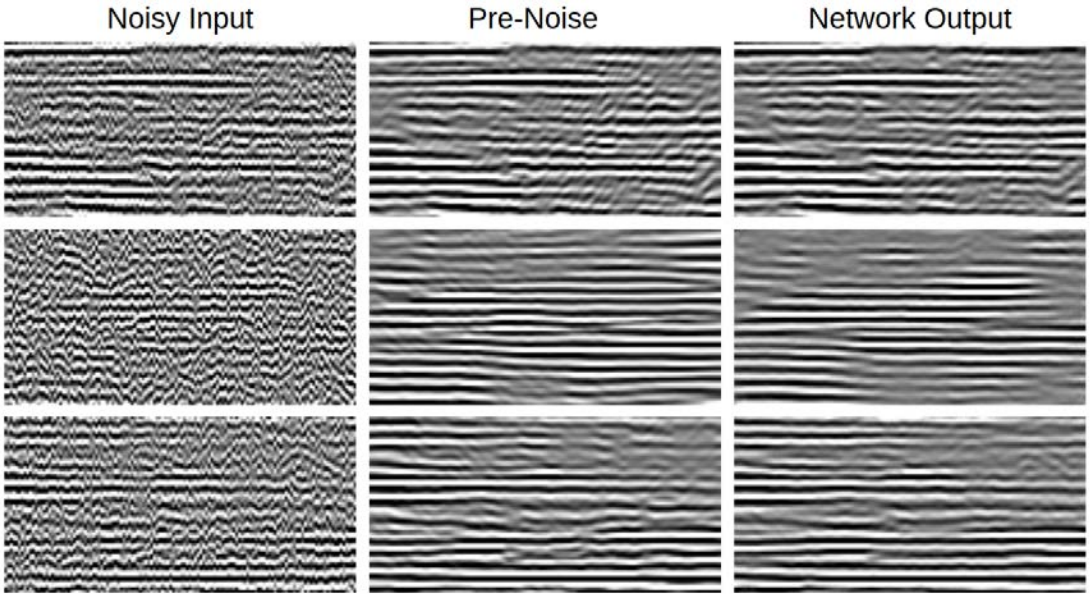

Fig. 3. Example network performance on validation set. Channel data of a single scan line are shown with the x-axis being the transducer element signals. Left column is the noisy simulated data, which is input into the network. The center column is the reference noise-free data. The right column is the output of the reverberation noise removal network.

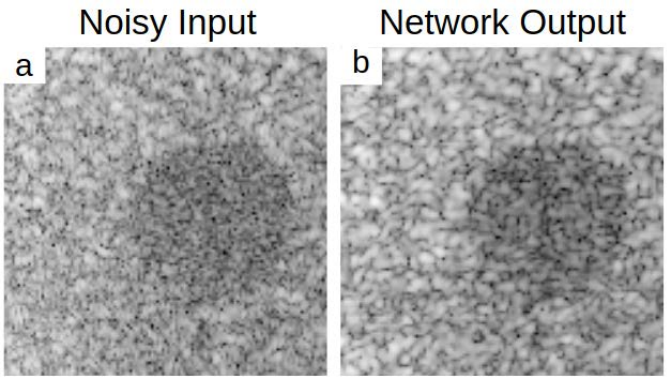

Fig. 4. B-mode images of phantom data with reverberation noise before and after being inferred on the network. a) was input to the network to give b). For this image pair, lesion contrast and CNR increased after network denoising, showing the network's ability to generalize to real data.

observation suggests that a comparison of signals between local transducer elements is important in identifying and removing reverberation noise.

Figure 3 shows example images of the RF channel signals with geometrical focal delays applied for a single scan line from the validation data before (left column) and after (right column) network inference along with their original, pre-noise simulated data as reference (center column). The noisy channel signals shown in Figure 3(left) are input to the 3D FCNN to give the output shown in Figure 3(right). The images show significant smoothing along the channels and closely resemble the noiseless data, indicating that the network successfully removed high-frequency reverberation noise in the channel dimension. Across the entire simulated validation set, the normalized RMS error decreased from 77% to 39% before and after inferring on the network, respectively. This further shows a large improvement of image quality on the simulated validation set after inferring on the reverberation removal network.

B-mode images generated by the neural network from the test phantom channel data are shown in Figure 4 before and after denoising on the network. Qualitatively, the reverberation noise inside the anechoic lesion was suppressed, as was the salt-and-pepper noise throughout the image, demonstrating the ability of the network to generalize to real data. Quantitatively, for the images in Figure 4 a and b, the lesion CNR was increased from 1.12 to 1.23 and the lesion contrast was increased from -9.9 to -12.2 dB.

Future work will focus on improving the generalization performance of the network on real-world data. This will be done by training the network with more realistic models of clutter. The current approach uses additive bandpass-filtered gaussian noise. Fullwave [16] simulations with and without reverberation noise will be used to further train the network towards generalizing to real data.

IV. Conclusion

In this work, a 3D fully convolutional neural network was trained on simulated data to remove reverberation noise. Training data was simulated using the Field II Pro simulation package, and reverberation clutter was approximated by adding bandpass filtered noise to the channel data. A 3D fully convolutional network was custom designed to encourage local comparison of lateral, axial and element data to remove reverberation clutter. The network was evaluated both in simulations and in experimental phantom data. The final

network decreased normalized RMS loss from 77% to 39% in the simulated test set, and qualitatively reduced noise and gave modest CNR improvements on real phantom data.

REFERENCES

[1] G. F. Pinton, G. E. Trahey, and J. J. Dahl, "Sources of image degradation in fundamental and harmonic ultrasound imaging: A nonlinear fullwave simulation study," vol. 58, no. 6, pp. 1272–1283, 2011.

[2] J. J. Dahl and N. M. Sheth, "Reverberation clutter from subcutaneous tissue layers: simulation and in vivo demonstrations," *Ultrasound in medicine & biology*, vol. 40, no. 4, pp. 714–726, 2014.

[3] M. A. Lediju, M. J. Pihl, J. J. Dahl, and G. E. Trahey, "Quantitative assessment of the magnitude, impact and spatial extent of ultrasonic clutter," *Ultrasonic imaging*, vol. 30, no. 3, pp. 151–168, 2008.

[4] S. P. Nasholm, R. Hansen, S. . Masoy, T. F. Johansen, and B. A. J. Angelsen, "Transmit beams adapted to reverberation noise suppression using dual-frequency surf imaging," *IEEE Transactions on Ultrasonics, Ferroelectrics, and Frequency Control*, vol. 56, no. 10, pp. 2124–2133, October 2009.

[5] "Ultrasonic reverberation clutter suppression using multiphase apodization with cross correlation," *IEEE Transactions on Ultrasonics, Ferroelectrics, and Frequency Control*, vol. 63, no. 11, pp. 1947–1956, Nov 2016.

[6] J. Shin and L. Huang, "Spatial prediction filtering of acoustic clutter and random noise in medical ultrasound imaging," *IEEE Transactions on Medical Imaging*, vol. 36, no. 2, pp. 396–406, Feb 2017.

[7] A. C. Luchies and B. C. Byram, "Deep Neural Networks for Ultrasound Beamforming," *IEEE Transactions on Medical Imaging*.

[8] J. A. Jensen and N. B. Svendsen, "Calculation of pressure fields from arbitrarily shaped, apodized, and excited ultrasound transducers," *IEEE Transactions on Ultrasonics, Ferroelectrics, and Frequency Control*, vol. 39, no. 2, pp. 262–267, 1992.

[9] J. A. Jensen, "Field: A Program for Simulating Ultrasound Systems," *Medical & Biological Engineering & Computing*, vol. 34, no. 1, pp. 351–353, 1996.

[10] ——, "A multi-threaded version of Field II," *IEEE International Ultrasonics Symposium, IUS*, pp. 2229–2232, 2014.

[11] O. Russakovsky, J. Deng, H. Su, J. Krause, S. Satheesh, S. Ma, Z. Huang, A. Karpathy, A. Khosla, M. Bernstein, A. C. Berg, and L. Fei-Fei, "ImageNet Large Scale Visual Recognition Challenge," *International Journal of Computer Vision*, vol. 115, no. 3, pp. 211–252, 2015.

[12] B. Zhou, A. Lapedriza, A. Khosla, A. Oliva, and A. Torralba, "Places: A 10 million Image Database for Scene Recognition," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 8828, no. c, pp. 1–14, 2017.

[13] X. Glorot, A. Bordes, and Y. Bengio, "Deep sparse rectifier neural networks," in *Proceedings of the Fourteenth International Conference on Artificial Intelligence and Statistics*, ser. Proceedings of Machine Learning Research, G. Gordon, D. Dunson, and M. Dudk, Eds., vol. 15. Fort Lauderdale, FL, USA: PMLR, 11–13 Apr 2011, pp. 315–323. [Online]. Available: http://proceedings.mlr.press/v15/glorot11a.html

[14] M. Abadi, P. Barham, J. Chen, Z. Chen, A. Davis, J. Dean, M. Devin, S. Ghemawat, G. Irving, M. Isard, M. Kudlur, J. Levenberg, R. Monga, S. Moore, D. G. Murray, B. Steiner, P. Tucker, V. Vasudevan, P. Warden, M. Wicke, Y. Yu, X. Zheng, and G. Brain, "TensorFlow: A System for Large-Scale Machine Learning TensorFlow: A system for large-scale machine learning," in *12th USENIX Symposium on Operating Systems Design and Implementation (OSDI '16)*, 2016.

[15] D. P. Kingma and J. Ba, "Adam: A method for stochastic optimization," *arXiv preprint arXiv:1412.6980*, 2014.

[16] G. F. Pinton, J. Dahl, S. Rosenzweig, and G. E. Trahey, "A heterogeneous nonlinear attenuating full-wave model of ultrasound." *IEEE Transactions on Ultrasonics, Ferroelectrics, and Frequency Control*, vol. 56, no. 3, pp. 474–88, 2009.