Performing Aperture Domain Model Image REconstruction (ADMIRE) Using a Deep Neural Network Sparse Encoder

Christopher Khan
Department of Biomedical Engineering
Vanderbilt University
Nashville, TN, USA
christopher.m.khan@vanderbilt.edu

Ruud J.G. van Sloun
Department of Electrical Engineering
Eindhoven University of Technology
Eindhoven, Netherlands
r.j.g.v.sloun@tue.nl

Brett Byram
Department of Biomedical Engineering
Vanderbilt University
Nashville, TN, USA
brett.c.byram@vanderbilt.edu

2021 IEEE International Ultrasonics Symposium (IUS) | 978-1-6654-0355-9/21/$31.00 ©2021 IEEE | DOI: 10.1109/IUS52206.2021.9593837

***Abstract*—Aperture Domain Model Image REconstruction (ADMIRE) is an advanced ultrasound beamforming method that uses a model-based approach to suppress sources of acoustic clutter and improve ultrasound image quality. However, although effective, ADMIRE requires solving an inverse problem that is ill-posed, which means that there are infinitely many solutions that can have different impacts on image quality. Currently, linear regression with elastic-net regularization is used to obtain a solution, but there are potentially better methods for performing model fitting. Therefore, in this work, we propose using a deep neural network sparse encoder for performing the model fits of ADMIRE. In particular, we unfold the iterations of the iterative shrinkage and thresholding algorithm (ISTA) as a feedforward neural network and train it using different training schemes to perform sparse coding. Test results using both simulated and *in vivo* data demonstrate that ADMIRE using a deep neural network sparse encoder has the potential to outperform conventional ADMIRE in terms of ultrasound image quality while still preserving the model-based intuition of ADMIRE.**

***Index Terms*—ADMIRE, Deep Learning, Sparse Coding, ISTA**

I. Introduction

Aperture Domain Model Image REconstruction (ADMIRE) is an adaptive beamformer that uses a model-based approach to suppress sources of acoustic clutter [1]–[3]. In particular, model matrices containing predictors representing the received aperture domain data, localized in time and frequency for different scattering locations, are fit to the ultrasound channel data. Once fit, the predictors within a region of interest (ROI) are used to reconstruct decluttered data. By using this process, we have demonstrated that ADMIRE can achieve significant improvements in image quality, and it can also do so in real time via an open-source, GPU-based implementation (https://github.com/VU-BEAM-Lab/ADMIRE) [1], [2], [4]–[8]. However, the model fitting process of ADMIRE requires solving an ill-posed, inverse problem. As a result, to obtain a solution, we currently use linear regression with elastic-net regularization, and we iteratively minimize the objective function in (1) using the cyclic coordinate descent optimization algorithm. In this equation, $\boldsymbol{X}$ is the ADMIRE model matrix, N is the number of observations per model predictor, P is the number of model predictors, $\boldsymbol{y}$ is the aperture domain frequency data to which the model matrix is fit, λ scales the amount of regularization, α is the weighting factor between L1-regularization and L2-regularization, and $\hat{\boldsymbol{\beta}}$ is the estimated model coefficients.

$$\hat{\boldsymbol{\beta}} = \arg\min_{\boldsymbol{\beta}} \frac{1}{2N} \sum_{i=1}^{N} \left(\boldsymbol{y}_i - \sum_{j=1}^{P} \boldsymbol{X}_{ij} \boldsymbol{\beta}_j \right)^2 + \lambda \left(\alpha \left\| \boldsymbol{\beta} \right\|_1 + \frac{(1-\alpha) \left\| \boldsymbol{\beta} \right\|_2^2}{2} \right) \tag{1}$$

Due to the ill-posed nature of this inverse problem, there are infinitely many solutions that can have different impacts on ultrasound image quality. The choice of using an elastic-net penalty provides just one of these solutions, but there are potentially better strategies for performing the model fits of ADMIRE. One of these possible strategies is to utilize deep learning, which has already demonstrated the ability to perform a variety of tasks in the field of ultrasound imaging using a data-driven approach [9]–[17]. Therefore, in this work we propose using a deep neural network sparse encoder in order to provide a data-driven approach for model fitting. Moreover, due to the fact that the neural network is only being used for model fitting instead of for end-to-end beamforming, the model-based intuition of ADMIRE is still preserved. A diagram of this approach is shown in Fig. 1.

II. Methods

In terms of the architecture of our network, we unfold the iterations of the iterative shrinkage and thresholding algorithm (ISTA) for sparse coding (i.e., solving $\boldsymbol{y} = \boldsymbol{X}\boldsymbol{\beta}$ for sparse code $\boldsymbol{\beta}$) as a feedforward neural network. This is based on previous work by Gregor and LeCun [18], who showed that this algorithm can be learned (LISTA) for performing sparse coding. The input to our neural network is $\boldsymbol{X}^\top \boldsymbol{y}$, and the output is $\hat{\boldsymbol{\beta}}$, which contains the predicted model coefficients for a given model fit. A diagram of the network architecture is shown in Fig. 2. In this diagram, **S** represents a fully connected

978-0-7381-1209-1/21/$31.00 ©2021 IEEE

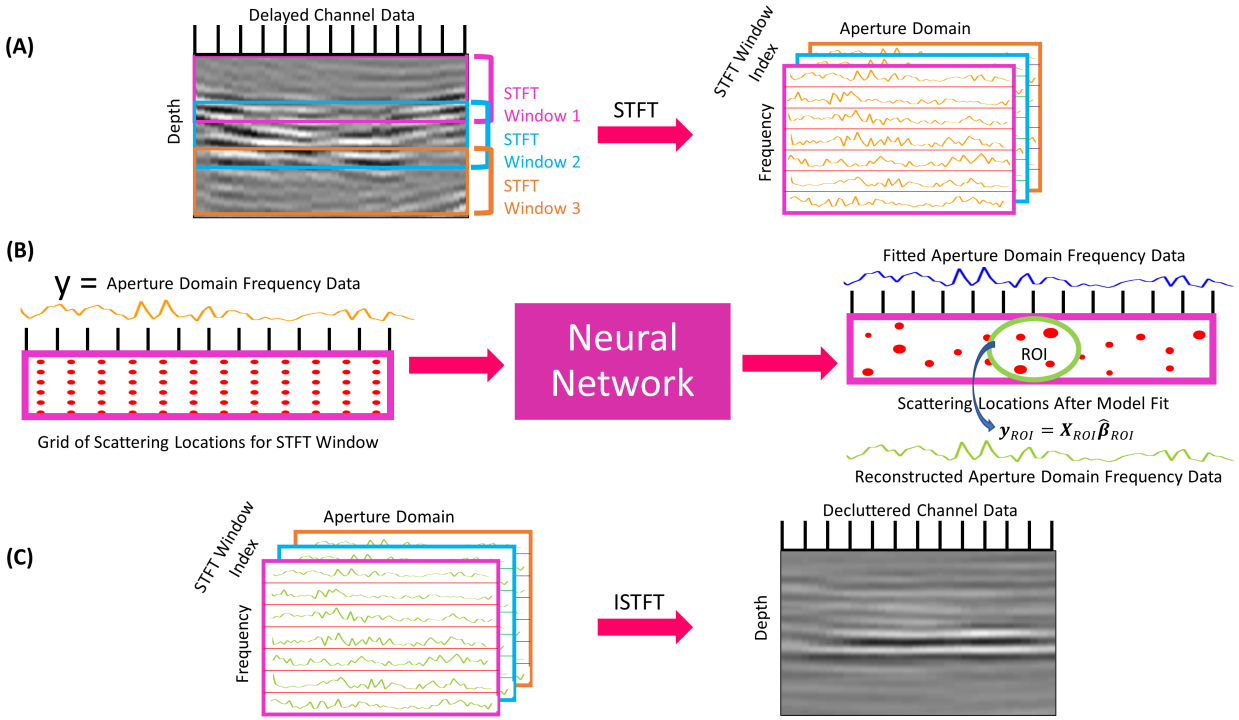

Fig. 1. Overview of ADMIRE. **(A)** Obtain the time-delayed channel data and calculate the short-time Fourier transform (STFT) along the depth dimension for each channel. **(B)** Obtain the corresponding model matrix for each set of aperture domain frequency data that will be reconstructed in each STFT window in each beam, fit each model matrix to its corresponding set of aperture domain frequency data using a deep neural network sparse encoder (sizes of red points correspond to how much each scattering location contributes to the aperture domain frequency data), and reconstruct each set of aperture domain frequency data by only using the predictors that correspond to scattering locations that are within a region of interest (ROI). **(C)** Calculate the inverse short-time Fourier transform (ISTFT) of the reconstructed aperture domain frequency data in order to obtain the decluttered channel data. Note that the scattering locations are not restricted to the depth range of the STFT window. The grid of scattering locations illustrated in **(B)** corresponds to the first STFT window. For STFT windows that correspond to deeper depths, the scattering locations can also be located in shallower depths because these locations can contribute to off-axis scattering and multipath scattering that affect the aperture domain frequency data for the STFT window. Essentially, as the depths become deeper for subsequent STFT windows, the depth range for possible scattering locations also increases.

linear layer, and a sigmoidal soft-thresholding function with a learnable threshold is used as the nonlinear activation function. The equation for this activation function is shown in equation (2), where x is the input to the function and γ is the learnable threshold.

$$\frac{x}{1 + e^{-(|x|-\gamma)}} \tag{2}$$

To train the network, data simulated in Field II [19], [20] consisting of hypoechoic and anechoic cysts with added reverberation clutter and noise was used, which allowed for the ground-truth clean and clutter signals to be known. The reverberation clutter was simulated using a pseudo nonlinear approach [21]. In addition, processing a single image frame with ADMIRE typically requires performing thousands of model fits, and each fit can be used as an individual training sample. In terms of the training schemes, we evaluated four distinct schemes. In the first scheme, we trained the network to simply reproduce the output of the ISTA algorithm. This was done by using a mean-squared error (MSE) loss function between the predicted model coefficients of the network for a given training iteration and those produced by the ISTA algorithm. In the other three schemes, the network was trained to learn its own fitting scheme rather than learning to mimic an existing scheme.

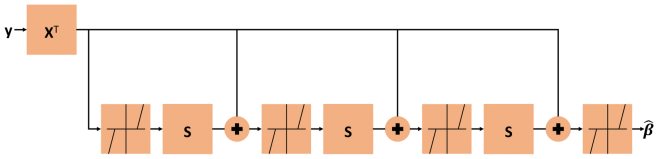

Fig. 2. Diagram of the deep neural network sparse encoder architecture for performing the model fits of ADMIRE.

In the second scheme, we trained the network to best reconstruct the clutter signal. This was done by taking the predicted model coefficients of the network for a given iteration, using them along with the corresponding ADMIRE model matrix to reconstruct the clutter signal as $\hat{\boldsymbol{y}}_{clutter} = \boldsymbol{X}_{clutter}\hat{\boldsymbol{\beta}}_{clutter}$, and then computing the MSE loss between the predicted clutter signal and the ground-truth clutter signal. In the third scheme, the same steps as the second scheme were used except that the network was trained to best reconstruct the clean signal instead of the clutter signal. Note that the clean signal refers to the signal without added reverberation clutter and noise. In

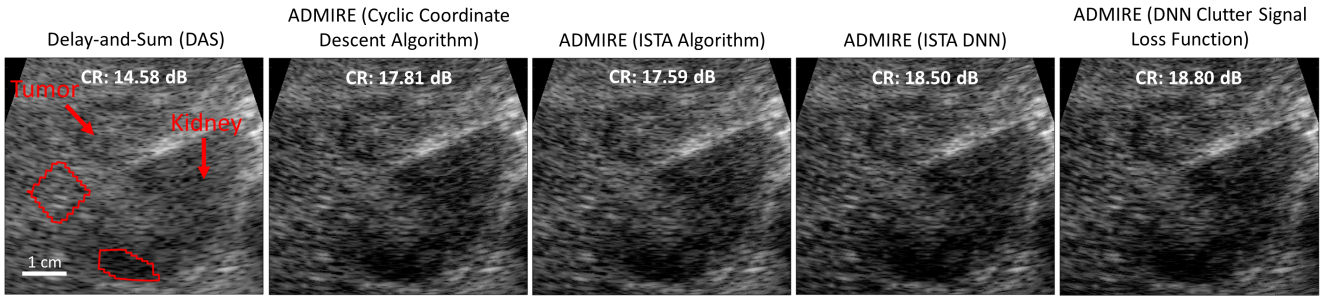

Image Metric	Ground-Truth Clean Signal	Delay-and-Sum (DAS)	ADMIRE (Cyclic Coordinate Descent Algorithm)	ADMIRE (ISTA Algorithm)	ADMIRE (ISTA DNN)	ADMIRE (DNN Clutter Signal Loss Function)
Contrast Ratio (dB)	25.75 ± 0.92	19.07 ± 1.02	23.53 ± 1.04	23.03 ± 1.06	24.06 ± 1.21	25.63 ± 1.33
Contrast-to-Noise Ratio (dB)	4.73 ± 0.70	4.16 ± 0.74	3.29 ± 0.91	3.19 ± 0.89	2.60 ± 0.82	3.17 ± 0.75
Generalized Contrast-to-Noise Ratio	0.98 ± 0.01	0.88 ± 0.04	0.92 ± 0.03	0.92 ± 0.03	0.92 ± 0.03	0.94 ± 0.03

Fig. 3. Top row: In vivo liver data images produced using different methods. For ADMIRE, the description in the parentheses for each case indicates which method was used to perform the model fits. The ISTA algorithm was used to perform linear regression with L1-regularization. For the cyclic coordinate descent algorithm case, linear regression with both L1-regularization and L2-regularization was performed with weighting factors of 0.9 and 0.1, respectively. All images are displayed with a dynamic range of 60 dB. The masks used for computing the contrast ratio values are displayed. Bottom row: Computed image metrics for Field II data sets (N = 5) processed with different methods. For purposes of training, both reverberation clutter and noise were considered to be clutter.

the fourth scheme, the network was trained to best reconstruct both the clutter and clean signals. This was done by computing the predicted clutter and predicted clean signals as well as the predicted combined signal given by $\hat{\boldsymbol{y}} = \boldsymbol{X}\hat{\boldsymbol{\beta}}$ for a given iteration and then using a joint loss function consisting of the addition of the clutter signal MSE loss, the clean signal MSE loss, and the combined signal MSE loss.

The PyTorch deep learning framework [22] was used for training, and the Adam optimizer [23] with a learning rate of 0.001 was utilized as the optimization algorithm. The same number of network layers as the network architecture shown in Fig. 2 was used, and the network weights were initialized to the same values between the training schemes by setting the pseudorandom number generator in PyTorch to the same initial seed. In addition, training was performed until either a maximum of 2,000 epochs were completed or if the validation loss function did not improve for 100 epochs. Once trained, these four methods were evaluated on image data from five test cyst simulations with added reverberation clutter and noise. In addition, other methods were evaluated, and these included delay-and-sum (DAS) beamforming, ADMIRE using cyclic coordinate descent to perform linear regression with elastic-net regularization ($\alpha = 0.9$), and ADMIRE using ISTA to perform linear regression with L1-regularization. Image metrics including contrast ratio, contrast-to-noise ratio, and generalized contrast-to-noise ratio were computed for each method and compared to the image metrics obtained using the ground-truth clean signal for the test cases. The methods were also evaluated on one frame of *in vivo* liver data, and contrast ratio values were computed. For both training of the neural networks and evaluation of the ADMIRE methods, fourth-order blind identification independent component analysis (FOBI-ICA) [7], [8] was applied in order to reduce the sizes of the ADMIRE model matrices.

III. Results

The test results for the different methods are shown in Fig. 3. Note that out of the three training schemes where the neural network learns its own model fitting scheme, the scheme in which the network was trained to best reconstruct the clutter signal had the best performance. Therefore, it is the only one out of the three that is shown. Moreover, for reconstructing the decluttered signal in ADMIRE, the predictors within an ROI are only used to directly reconstruct it due to the fact that the predictors outside of the ROI contribute to clutter. However, for the network that was trained to best reconstruct the clutter signal, we found that the performance was further improved by using the predictors outside of the ROI to reconstruct the clutter signal and then subtracting it from the aperture domain frequency data to obtain the decluttered signal. This is similar to the approach used for iterative ADMIRE [24]. In addition, for the neural network trained to mimic ISTA, it could be trained using either simulated or *in vivo* data due to the fact that the ISTA algorithm can be applied to both. As a result, it was trained two separate times. In one case, it was trained

with only the simulated cyst data, and in another case, it was trained with only one separate frame of *in vivo* liver data. The network had better performance when trained with *in vivo* data, so that is the one that is shown.

IV. Discussion

In terms of the test cyst simulation results, using ADMIRE with the deep neural network that was trained to best reconstruct the clutter signal produced image metrics that had the overall greatest similarity to the image metrics for the ground-truth clean signal. While it is possible to change the regularization parameters of conventional ADMIRE (i.e., no neural network), the ground-truth clean signal is not known *in vivo*, which means that a user would not know which parameter values to use to get closer to it. Therefore, we utilized empirically determined parameter values that we typically use for processing with ADMIRE instead of attempting to tune them. In addition, for the *in vivo* liver data test case, the network that was trained to best reconstruct the clutter signal also produced the highest contrast ratio value. Therefore, these results show that a deep neural network sparse encoder can be trained to perform the model fits of ADMIRE. Moreover, they also demonstrate that using ADMIRE with a deep neural network sparse encoder has the potential to outperform conventional ADMIRE in terms of ultrasound image quality while still preserving its model-based intuition.

V. Conclusion

In this work, we demonstrated that a deep neural network sparse encoder can be trained to perform the model fits of ADMIRE, and we also showed that this data-driven approach has the potential to outperform conventional ADMIRE in terms of ultrasound image quality. Moreover, due to the fact that the neural network in this approach is only determining an optimal linear combination of features that we construct with our own physics-based model, we still maintain the interpretability and intuition of ADMIRE that would be lost if an end-to-end beamforming approach were to be used.

VI. Acknowledgement

The authors would like to thank the staff of the ACCRE computing resource. This work was supported by NIH grants R01HL156034, R01EB020040, and S10OD016216-01, NAVSEA grant N0002419C4302, and NSF award IIS-1750994.

References

[1] B. Byram and M. Jakovljevic, "Ultrasonic multipath and beamforming clutter reduction: A chirp model approach," *IEEE Trans. Ultrason., Ferroelectr., Freq. Control*, vol. 61, no. 3, pp. 428–440, 2014.

[2] B. Byram, K. Dei, J. Tierney, and D. Dumont, "A model and regularization scheme for ultrasonic beamforming clutter reduction," *IEEE Trans. Ultrason., Ferroelectr., Freq. Control*, vol. 62, no. 11, pp. 1913–1927, 2015.

[3] K. Dei and B. Byram, "The impact of model-based clutter suppression on cluttered, aberrated wavefronts," *IEEE Trans. Ultrason., Ferroelectr., Freq. Control*, vol. 64, no. 10, pp. 1450–1464, 2017.

[4] C. Khan, K. Dei, S. Schlunk, K. Ozgun, and B. Byram, "A Real-Time, GPU-Based Implementation of Aperture Domain Model Image REconstruction." *IEEE Trans. Ultrason., Ferroelectr., Freq. Control*, vol. 68, no. 6, pp. 2101–2116, Feb. 2021.

[5] C. Khan and B. Byram, "GENRE (GPU Elastic-Net REgression): A CUDA-accelerated package for massively parallel linear regression with elastic-net regularization," *J. Open Source Softw.*, vol. 5, no. 54, p. 2664, 2020.

[6] K. Dei, S. Schlunk, and B. Byram, "Computationally efficient implementation of aperture domain model image reconstruction," *IEEE Trans. Ultrason., Ferroelectr., Freq. Control*, vol. 66, no. 10, pp. 1546–1559, Oct. 2019.

[7] J. F. Cardoso, "Source separation using higher order moments," in *Proc. Int. Conf. Acoust., Speech, Signal Process.*, 1989, pp. 2109–2112.

[8] J. Shlens, "A tutorial on independent component analysis," 2014, *arXiv:1404.2986*. [Online]. Available: http://arxiv.org/abs/1404.2986

[9] R. J. Van Sloun, R. Cohen, and Y. C. Eldar, "Deep learning in ultrasound imaging," *Proceedings of the IEEE*, vol. 108, no. 1, pp. 11–29, 2019.

[10] A. C. Luchies and B. C. Byram, "Deep neural networks for ultrasound beamforming," *IEEE Trans. Med. Imag.*, vol. 37, no. 9, pp. 2010–2021, 2018.

[11] J. Tierney, A. Luchies, M. Berger and B. Byram, "Evaluating Input Domain and Model Selection for Deep Network Ultrasound Beamforming," *IEEE Trans. Ultrason., Ferroelectr., Freq. Control*, vol. 68, no. 7, pp. 2370-2385, Jul. 2021.

[12] J. Tierney et al., "Training Deep Network Ultrasound Beamformers with Unlabeled In Vivo Data," *IEEE Trans. Med. Imag.*, doi: 10.1109/TMI.2021.3107198.

[13] L. L. Brickson, D. Hyun, M. Jakovljevic, and J. J. Dahl, "Reverberation noise suppression in ultrasound channel signals using a 3d fully convolutional neural network," *IEEE Trans. Med. Imag.*, vol. 40, no. 4, pp. 1184–1195, 2021.

[14] B. Luijten et al., "Adaptive ultrasound beamforming using deep learning," *IEEE Trans. Med. Imag.*, vol. 39, no. 12, pp. 3967–3978, 2020.

[15] A. Wiacek, E. Gonz´alez, and M. A. L. Bell, "Coherenet: A deep learning architecture for ultrasound spatial correlation estimation and coherence-based beamforming," *IEEE Trans. Ultrason., Ferroelectr., Freq. Control*, vol. 67, no. 12, pp. 2574–2583, 2020.

[16] D. Hyun, L. L. Brickson, K. T. Looby, and J. J. Dahl, "Beamforming and speckle reduction using neural networks," *IEEE Trans. Ultrason., Ferroelectr., Freq. Control*, vol. 66, no. 5, pp. 898–910, 2019.

[17] R. J. G. van Sloun et al., "Super-Resolution Ultrasound Localization Microscopy Through Deep Learning," *IEEE Trans. Med. Imag.*, vol. 40, no. 3, pp. 829-839, March 2021.

[18] K. Gregor and Y. LeCun, "Learning fast approximations of sparse coding," in textitProceedings of the 27th international conference on international conference on machine learning, 2010, pp. 399-406.

[19] J. A. Jensen, "FIELD: A program for simulating ultrasound systems," in *Proc. 10th Nordic-Baltic Conf. Biomed. Imag.*, 1996, vol. 4, no. 1, pp. 351–353.

[20] J. A. Jensen and N. B. Svendsen, "Calculation of pressure fields from arbitrarily shaped, apodized, and excited ultrasound transducers," *IEEE Trans. Ultrason., Ferroelectr., Freq. Control*, vol. 39, no. 2, pp. 262–267, Mar. 1992.

[21] B. Byram and J. Shu, "A pseudo non-linear method for fast simulations of ultrasonic reverberation," in *Medical Imaging 2016: Ultrasonic Imaging and Tomography*, International Society for Optics and Photonics, vol. 9790, 2016, 97900U.

[22] A. Paszke et al., "Pytorch: An imperative style, high-performance deep learning library," in *Advances in Neural Information Processing Systems 32*, H. Wallach, H. Larochelle, A. Beygelzimer, F. d'Alch´e-Buc, E. Fox, and R. Garnett, Eds., Curran Associates, Inc., 2019, pp. 8024–8035.

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

[24] S. Schlunk, K. Dei, and B. Byram, "Iterative Model-Based Beamforming for High Dynamic Range Applications," *IEEE Trans. Ultrason., Ferroelectr., Freq. Control*, vol. 68, no. 3, pp. 482-493, Mar. 2021.