Learning to Reconstruct Speckle-Free Ultrasound B-mode Images

Leandra Brickson

PI: Jeremy Dahl

Abstract

Ultrasound B-mode images are maps of the echogenicity of the imaged media. With traditional beamforming methods, B-mode images contain speckle noise caused by the random interference of sub-resolution scatterers . In this poster, I present a framework for posing ultrasound beamforming and image reconstruction as a machine learning problem and utilize 3D fully convolutional encoding- decoding networks to learn a mapping from raw data from the ultrasound scanner to images of echogenicity.

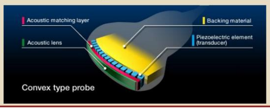
Project Goal

Train a fully convolutional neural networks to recover a de-speckled echogenicity (sound reflectivity) map from raw ultrasound image channel data.

Speckled Ultrasound Image — Trained Neural Network — Speckle – Free Echogenicity Map

Datasets & Simulations

To train the network, the Field II Pro simulation package [1] was used to simulate ultrasound channel data from the ImageNet and Places datasets. The training set contained 100k 32x32 pixel patches of beamformed images and 20k of 32x32x16 pixel patches of channel signal images. Ongoing work includes continuing to expand this dataset to expose the network to as wide a variety of imaging conditions as possible.



The simulated data provides the raw channel data from each transducer element. For network evaluation, simulated, and full synthetic aperture phantom and in-vivo data were used.

Optimization

L2 and L1 loss functions were used for optimization, Total Variation [4] and Structural Similarity[3] were also tested, but didn't provide as good results.

L2 Difference Loss

$$\sum_{i=0}^{n}(\hat{y}_i - y_i)^2$$

Structural Similarity

$$\frac{(2\mu_{\hat{y}}\mu_y + c_1)(2\sigma_{\hat{y}y} + c_1)}{(\mu_{\hat{y}}^2 + \mu_y^2 + c_1)(\sigma_{\hat{y}}^2 + \sigma_y^2 + c_2)}$$

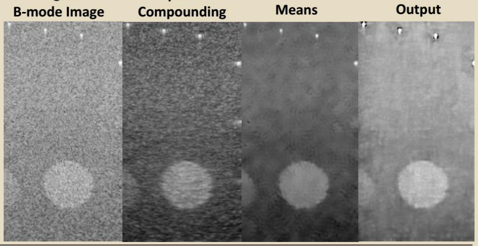
Total Variation Prior

$$\sum_{i,j}\sqrt{\left|\hat{y}_{i+1,j} - \hat{y}_{i,j}\right|^2 + \left|\hat{y}_{i,j+1} - \hat{y}_{i,j}\right|^2}$$

The network was optimized using the Adam optimizer and the hyperparameters tuned were L1 and L2 regularization, learning rate, TV loss regularization, Log loss(binary), batch size, dropout, number of filters, and network depth

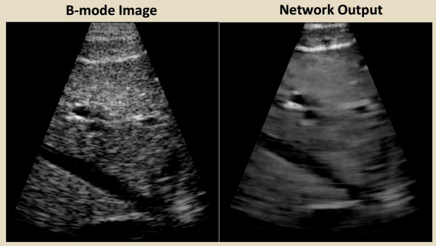
Results

Phantom



	Original B-mode Image	Spatial Compounding	Non-Local Means	Network Output
SNR	1.91	3.14	15.4	9.34
CNR	0.57	0.62	0.752	3.3

In-Vivo

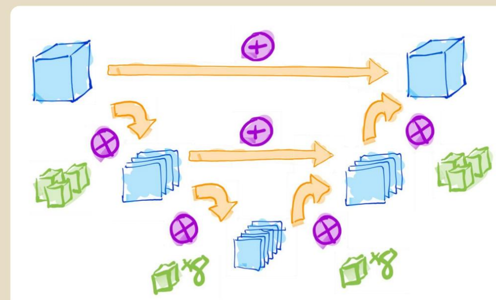


Original B-mode Image — Primitive Network Output

While the network works well on phantom data, it is yet to be seen it if works well on In-Vivo data, due to limitations on the test dataset.

Network

The network architecture is inspired by U-Net. It is a 3D fully-convolutional encoding-decoding network. Down-sampling is done by performing a 3D convolution , and then max-pooling in 2D. Up sampling is done by performing a 3D transpose convolution and then concatenating with previous layers in the encoding layer.



References:

[1] Jensen, Jorgen Arendt and Svendsen, Niels Bruun (1992). 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.*

[2] Abbott, John G. and Thurstone, F. L., Acoustic speckle: Theory and experimental analysis. *Ultrasonic Imaging*

[3] Zhou Wang ; A.C. Bovik ; H.R. Sheikh ; E.P. Simoncelli , Image quality assessment: from error visibility to structural similarity, IEEE Transactions on Image Processing (Volume: 13, Issue: 4, April 2004)

[4] C.R. Vogel ; M.E. Oman, Fast, robust total variation-based reconstruction of noisy, blurred images, *IEEE Transactions on Image Processing* (Volume: 7, Issue: 6, Jun 1998)

[5] Zhou, Bolei and Lapedriza, Agata and Khosla, Aditya and Oliva, Aude and Torralba, Antonio, Places: A 10 million Image Database for Scene Recognition, vol. 8828, no. c, pp. 1-14, 2017.